THE PROCEEDINGS OF
THE PHYSICAL SOCIETY

VOL. 46, PART 4 *July* 1, 1934 NO. 255

551·593·41

# FURTHER EXPERIMENTS IN ILLUSTRATION OF THE GREEN FLASH AT SUNSET

BY LORD RAYLEIGH, F.R.S.

*Received March* 21, 1934. *Read April* 20, 1934.

*ABSTRACT* An experimental imitation of the phenomenon of

---

## ERRATUM

*Proc. Phys. Soc.* **46**, 371, 378 (1934).

Transpose the words

(*Discussion continued on page* 478)

from page 371 to page 383.

---

appearance at the vanishing limb of the sun, when sunset occurs under a sea, desert, or mountain horizon*. It has of recent years been usually attributed to atmospheric dispersion, which would naturally keep the more refrangible part of the spectrum in view when the rest had vanished.

In a paper published in 1930† I discussed whether the normal atmospheric dispersion, calculated from the known refraction at the horizon (35′) and the known dispersion of air, was adequate in amount to produce this effect, and concluded that it *was* adequate. The essential point of the discussion is that the length of spectrum produced by the atmospheric prism varies as the direct distance of the horizon from the observer, whereas the linear size of the diffraction pattern, defining the diffuseness of the shadow, varies as the square root of the distance. Hence with increasing distance the conditions become increasingly favourable for observing a definite chromatic border at the edge of the shadow; at a distance of several miles the shadow is more than sharp enough to give the chromatic effects full scope. In other words the atmospheric spectroscope has adequate resolving power, its effective aperture being, not that of the eye, but a much larger aperture which may be considered as located at the horizon.

---

* General references are: A. A. Rambaut, *Symons' Meteorological Magazine*, **41**, 23, 40 (1906); M. E. Mulder, *The Green Ray*, London, 1922; P. Feenstra Kuiper, *De Groene Strahl*, 1926.

† *Proc. R.S.* **126**, 311 (1930).

These conclusions were illustrated by means of a crown glass prism of angle 45′ which, it was calculated, had the same dispersive power as the atmosphere near the horizon. A straight edge representing the horizon was placed across the aperture at the thick end and parallel to the base. When a beam of sunlight was admitted through this prism and received on a screen 20 metres away, the shadow cast by the edge was seen to be distinctly bordered with colour. The colour was described as greenish blue.

The present paper describes additional experiments intended to illustrate the green flash. They deal first with the imitation of the effect by means of a prism having dispersion equal to that of the atmosphere without special reference to the exact colour; and secondly with the imitation of those conditions which usually determine the colour to be green rather than blue.

## § 2. A MODIFIED EXPERIMENT IN IMITATION OF THE GREEN FLASH

The experiment described in the former paper and in the introduction, though conclusive enough as a verification, differs somewhat from the natural green flash in that the colour is observed on a screen rather than received direct into the eye. The direct light of the sun when high up is too bright for convenient observation, and moreover the fringe of colour at the available distance of 20 metres from the obscuring edge is too narrow to fill the pupil of the eye, and it becomes lost in the general glare. I have sought to modify the experiment so as to imitate more closely the natural effect.

To do this it is necessary to use longer distances than are available in the laboratory, and therefore to work out of doors. To avoid the excessive brightness of the high sun an artificial source was used. Since this is to be observed at a considerable distance it should have large linear dimensions, say $\frac{1}{100}$ of the distance from the observer, as is roughly the case with the sun.

A strip of opal glass, backed by a row of nine incandescent lamps, 25 watts each, was used. The lamps were in a box, with the opal strip let into the side. The strip was masked so as to present a luminous segment of about 3·4 metres radius, and to minimize false light it was eventually found necessary to supplement the mask by a second mask, so that the visible part of the opal was a circular arc*, 2 mm. broad, figure 1. The prism used to imitate atmospheric dispersion requires to have a long refracting edge, say half the length of the source or 50 cm. long. To avoid the expense and delay of having a special glass prism made, a water prism was assembled from plate-glass strips, measuring 50 × 7 cm. and 1 cm. thick. These were placed together at an acute angle with two of the long edges in contact and with slips of window glass to cover the side and bottom edges. The whole was assembled with Prout's elastic glue. This cement becomes viscous in a hot room, and it was possible to adjust the angle by the use of gentle mechanical force without taking the prism to pieces.

---

* A narrow segment would represent the vanishing limb of the sun better than an arc of light, but in other respects is less convenient. The point is of secondary importance.

The refraction of the atmosphere at the horizontal is 35′ and the relative dispersion $(\mu_F - \mu_C)/(\mu_D - 1)$ is $1{\cdot}04 \times 10^{-2}$. In order that a water prism (with relative dispersion $1{\cdot}8 \times 10^{-2}$) should be equivalent to this, its deviation should be $35' \times 1{\cdot}04/1{\cdot}8$ or 20′. The water prism as set up gave a measured deviation of 21′, which is a satisfactory approximation. Since the thick end of the prism was open, this end was necessarily uppermost, so that the whole arrangement was inverted as compared with the atmospheric refracting layer, and the light passed below the horizontal obscuring edge representing the horizon. This edge was, of course, placed so as to cut off the thick end of the prism. The source was 56 metres on one side of the prism, and the observer 74 metres on the other. When the eye was got into the right position, the source could be made to vanish behind the obscuring edge by a slight upward movement of the head and the greenish-blue disappearing flash could then be seen; or, by keeping the eye in an appropriate position, it could be held in view. If an ocular diaphragm was placed in the most favourable position, so that the colour was seen at the bottom (horizontal) position of the disappearing arc, then the colour at the bottom was lost in darkness if the diaphragm was raised 1 cm., and was lost in full illumination if it was lowered 1 cm. I myself was able to observe best with a Galilean telescope giving × 2 magnification. My assistant, Mr R. Thompson, whose vision is unimpaired, preferred the naked eye*.

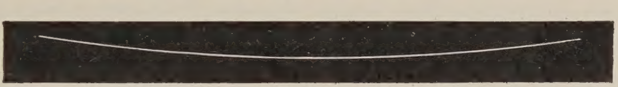


Figure 1.

If the eye is lowered slightly from the position where the colour is seen at the bottom of the arc, the colour retreats towards the higher parts of the source on either side, the bottom being yellow or white by comparison. The two stages are roughly shown in figure 2, the shading indicating green. This effect is not so well seen with a luminous segment, like the setting sun, as with a luminous arc. In that case the glare due to the fully exposed segment is too great. In the present experiment it was found that to work with a full segment made the observations much less satisfactory, the aperture of the prism being filled with false (scattered) light when viewed from inside the shadow. This circumstance is instructive in connexion

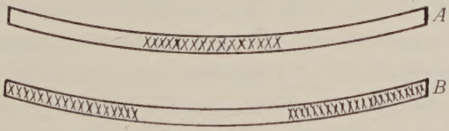


Figure 2.

---

* A small telescope is generally found advantageous for observing the natural green flash at sunset.

with the apparent caprice in the visibility of the natural green flash. A little false light due to atmospheric agencies is enough to obliterate it.

The experiment described seems to approach as nearly as could be wished to an artificial reproduction of the green flash, the question of the exact colour being left aside for a moment. Since the distances used are so very much smaller than those in the natural case, there appears to be no special point in reducing them still further except as a matter of practical convenience. It may be mentioned, however, that the colour could still be well seen when the distance from the prism to the observer was reduced to 36 metres.

The source limited to an arc 2 mm. wide, as described, may be considered to approximate to a spectroscopic slit, and on a first view of the matter it might be expected that a spectrum would be seen on looking at it through the clear aperture of the prism: or at any rate that one side would be visibly bordered by a blue-green fringe, and the other by a red. Such is not the case, and it is only when the obscuring edge at the prism is brought into action that the blue-green colour is seen. This circumstance illustrates very definitely the view briefly explained in the introduction, and more fully in the paper of 1930, that the adequacy of the resolving power essentially depends on the action of the horizon.

## § 3. THE CHROMATIC QUALITY OF THE FLASH

The view that the appearance of the green flash is due in one way or another to physiological contrast is an old one. The most obvious application of this idea is that it is merely the result of retinal fatigue. After the more or less reddish disc of the sun has been looked at, retinal fatigue shows a green appearance when this has vanished. The particular explanation called "successive contrast" is ruled out by the fact that the green flash appears at sunrise. Dr J. S. Haldane has, however, recently revived the same point of view in a different form*. He says: "Lord Rayleigh devised an experimental arrangement which enabled the effect to be studied at leisure. The colour seen was, however, blue, not green. It now seems evident, in view of the experiments described above, that the cause of the actual sunset or sunrise coloured ray being vividly green, is the simultaneous contrast effect due to the light coming from round the sun being red."

Dr Haldane has kindly informed me, however, in reply to a question, that like most recent writers, including myself, he regards atmospheric dispersion as the primary cause of the green flash. In his view simultaneous contrast comes in as a secondary cause to modify the subjective colour effect.

It is desirable to avoid any over-emphasis of the point to be explained. Although the colour is usually described as some variety of green such as bluish green, emerald green, or grass green, there are a number of records in the literature where blue is mentioned†. However, it is a fact that in most of the records green alone is mentioned.

---

* *J. Physiol.* **79**, 129 (1933); *The Times*, June 10, 1933.

† For example: Lord Kelvin, *Nature*, **60**, 411 (1899); E. H. Hills, *Monthly Note R.A.S.* **62**, 430 (1902); Sir Flinders Petrie, *The Times*, June 3, 1933. In the latter two publications the blue phase is described as following the green at sunset or preceding it at sunrise.

We are not here concerned with the more general aspect of Haldane's investigation of the physiological significance of colour contrast, but I have repeated those of his experiments which seem to be most in point, and have found them interesting and instructive. The clearest form of experiment is to compare the effect of two coloured lights on a sheet of foolscap paper surrounded by black velvet; one half of the paper is illuminated by the first source and the other by the second source. The arrangement is simply that of Rumford's photometer: the black velvet eliminates any complication due to outer parts of the screen being illuminated by both sources. The most striking effects are obtained when the lights are of comparable intensity; this is secured by increasing the distance of the brighter light from the screen.

Dr Haldane finds that if one side is illuminated by a candle and the other by a red lamp, the former will appear a "deep green." I have myself been more successful in producing a green by using a daylight lamp as one source, and a red or orange lamp as the other. The daylight lamp gives white when used alone, but by contrast with the red or orange it appears what I should call a bluish green. This is the best success I have been able to get in producing by contrast a green from light that was not green by itself. I could not call it bright green, or grass green, or emerald green.

But, after all, the really relevant question is whether the blue of the spectrum can be made to look green by being contrasted with a red or orange light. To test this by Dr Haldane's methods, I used the blue radiation from a mercury vapour lamp with a suitable colour-filter to eliminate the green and yellow mercury lines, leaving the blue line $\lambda 4358$ virtually isolated. The blue patch from this was contrasted with a variety of orange and red sources with varied intensity, but neither my assistant Mr R. Thompson nor myself could describe it as anything but blue throughout*.

The above experiments did not give much reason to hope that the bluish flash could be turned into a green one by any method depending on colour contrast. But before abandoning the idea it was thought best to resort again to the direct method of experiment, prismatic dispersion being used to imitate atmospheric dispersion.

In experimenting on the purely chromatic aspect of the problem it is evidently unnecessary to use the rather arduous conditions of observation described in § 2. Ample prismatic dispersion may be used, which allows of short distances and a field of colour large enough for easy and convenient study in the laboratory. I have used a crown glass prism of angle 24°. This was set up in the same general manner, but the refracting edge was vertical, and, as before, an obscuring edge was placed over the thick end. An incandescent lamp with opal bulb served to represent the sun. The observer was 6 metres from the lamp with the prism half way. The general arrangement is shown in figure 3 (not drawn to scale). The disappearing flash seen with this arrangement as the eye moved into the shadow was of course dark blue or violet. The problem was to find what modification of conditions would make it a bright green.

---

* I do not enter upon these more or less physiological questions without misgiving, bred partly of lack of familiarity with this field, and partly from dislike of the vagueness inherent in the use of words for describing subjective colour impression. It is certain that some people describe as blue what others describe as green.

I first tried whether a background of orange or red would produce this effect, as Dr Haldane has supposed might be the case. Various samples of red, orange and yellow cloth or paper were placed so as to be strongly illuminated by the lamp, and so as to form a background on which the disappearing flash could be seen. These coloured backgrounds could be alternated with white paper or with black velvet. The white or coloured background of course tended to make the blue flash less pure, but I could not by these methods get any approximation to green. The brightness of the background by this method was not under very convenient control.

An alternative method was to observe the disappearing flash against a self-luminous orange or red background formed by a second opal lamp with a coloured glass superposed. It was important so to tilt the coloured glass that the specular reflection of the main lamp in it was not visible at the point of observation. By controlling the voltage at the terminals of the second lamp the brightness of the coloured background could be controlled; but again I was not able to get anything but a blue disappearing flash.

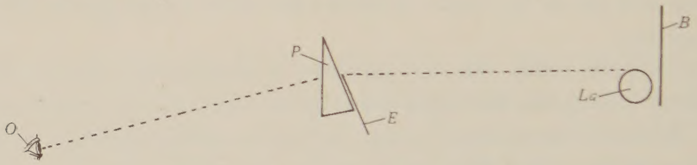


Figure 3. *O*, observer; *P*, prism; *E*, observing edge; *La*, lamp; *B*, coloured background.

Instead of coloured glass we may use as background the orange light transmitted through a scattering medium (rosin in water—see below) which approximates closely in physical causation to the orange colour of the sunset sky. This background was tried at various intensities, but again the flash remained blue and not green.

Thus in spite of persistent efforts no success has been obtained in reproducing the green colour by simultaneous contrast.

It has often been suggested that the green colour of the residual flash is to be explained by atmospheric scattering. Light from the low sun has to traverse a thick atmospheric layer, which preferentially scatters blue light laterally and transmits light of lower refrangibility, thus acting in effect as a yellow or orange filter which removes blue and leaves green as the most refrangible component of the spectrum. The scattering power of the atmosphere is variable according to the number and size of the dust particles present, and here we have the flexibility of conditions which is necessary to explain why the flash is usually described as green or blue-green, but occasionally as blue.

Pursuing the policy of experimental illustration, I have used a liquid scattering-medium in front of the source to represent the atmosphere. A metal tank 90 cm. long with glass ends was provided, large enough to transmit a beam of diameter 8 cm. This was set up next the source, between it and the prism, figure 4. It was filled with 7 litres of water, and a solution containing 1 gm. of rosin in 100 cm$^3$ of alcohol was progressively added and well stirred in. This method of making a suspension of fine particles for scattering light is a very old one, originally due to Brücke*.

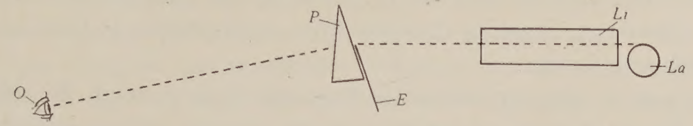


Figure 4. *O*, observer; *P*, prism; *E*, scattering edge; *Li*, scattering liquid; *La*, lamp.

With the gas-filled opal lamp at full voltage, the disappearing flash was observed. The results were as shown in the table.

| Total quantity of solution added (cm$^3$) | Colour of lamp viewed through liquid without prism | Colour of disappearing flash |
|---|---|---|
| 5 | Pale yellow | Blue |
| 10 | Yellow | Bluish green |
| 15 | Pale orange | Full green (optimum) |
| 20 | Orange | Green, but diminished in intensity |

The action of a scattering medium is thus well able to produce a full green colour in the disappearing flash.

The amount of scattering required is such as to make the source appear a pale orange. The setting sun is, of course, often seen blood red, so that the preferential scattering of blue light postulated is by no means excessive. Probably the green flash would not be seen with a blood-red sun; but so far as one can judge without the opportunity of observing in tropical seas, it would seem that this pale orange about represents the colour of the setting sun under conditions favourable for observing the green flash.

## DISCUSSION

Instructor-Capt. T. Y. Baker. The author has stated in this paper, and, I understand, in his previous paper of 1930, that the normal atmospheric dispersion calculated from the known refraction at the horizon (35′) and the known dispersion of air, is adequate to produce the phenomenon of the green flash. If I understand his explanation correctly, it is based upon figures giving the refractive index of air for two colours, red and green. The values measured by Meggers and Peters† are: red 1·00029088, green 1·00029392. The dispersion of air, calculated like the dispersion of glass, gives a $\nu$ value of 96 just as a crown glass has a $\nu$ value of 60 and a flint glass one of 36. One could make a crown-glass prism with a deviation of 35′, when the dispersion between red and green would be 35′/60 or 35″; or a flint prism of the same deviation, when the dispersion would be 35′/36 or 58″. If one could make an air prism to give this deviation the dispersion would be 35′/96 or 22″, which corresponds roughly to some of the observed durations of the green flash when it has been seen. I gather that it is upon a calculation of this nature that the author bases his statement.

---

* *Pogg. Ann.* **88**, 382 (1853). Brücke actually used mastic.

† Bureau of Standards, No. 327.

My objection to this explanation is that refraction through the atmosphere is different in its nature from refraction through a prism, and an explanation based upon measurements of refractive index and dispersion at sea-level densities alone is not adequate. Meggers and Peters's figures for red and green may possibly be questioned with reference to the last decimal place, but no one will dispute that the difference of three units in the sixth place does substantially represent the difference of refractive index for red and green light. It is a figure that is valid at any place on the earth's surface at any time. That being so, if the author's explanation is adequate the green flash should always be seen when the sun sets cleanly behind the horizon. Naturally one would expect small variations of the 22″, but not variations amounting to complete elimination of the dispersion. And it must be remembered that the green flash is not a common phenomenon. It is the exception rather than the rule. From personal experience, although I have watched numerous sunsets and a few sunrises from the deck of a ship in the Mediterranean, I have never seen it there, and my first and only experience of it was at Southsea last year. On that occasion the duration was probably not more than $\frac{1}{5}$ sec., corresponding to 4″ or 5″.

The true explanation of the green flash is, in my opinion, much more complicated than the author suggests. Any consideration of astronomical refraction must necessarily be based on the assumption that the atmosphere is stratified in spherical layers concentric with the earth. In this non-homogeneous medium the path of a ray can be expressed in the form $np =$ constant. That formula is to be found in many text-books, but perhaps the most appropriate reference I can give is Lord Rayleigh, *Phil. Mag.* **36**, 141 (1893). In the formula $n$ is the refractive index at any point of the medium and $p$ the length of the perpendicular from the centre of the earth upon the tangent line to the ray at the point in question. The stratification being spherical, $n$ is a function of height only, that is to say a function of $r$, the distance of the point from the centre of the earth. From the equation it is a simple matter to derive the value of $\psi$, the refraction of a ray which, entering the atmosphere from the outside, becomes tangential to the earth at sea level. Leaving out of account the additional refraction between the contact point and the observer's eye, the amount of the refraction up to the contact point is

$$\psi = \int_{r_0}^{R} \frac{n_0 r_0 \,(dn/dr)\, dr}{n \sqrt{(n^2 r^2 - n_0{}^2 r_0{}^2)}},$$

where $n_0$ is the refractive index at the earth's surface, $r_0$ the radius of the earth, and $R$ the radius to the extreme limit of the atmosphere. The phenomenon of the green flash requires that integral to have a larger value for green light than for red. If one accepts 22″ as being the average dispersion corresponding to the observed durations of the flash, then $\psi_{\text{green}} - \psi_{\text{red}} = 22''$. But it seems to me evident that

$\psi_{\text{green}} - \psi_{\text{red}}$ depends not only on $n_{0\,\text{green}}$ and $n_{0\,\text{red}}$, which is all that the author has taken into account, but on $n_{\text{green}}$ and $n_{\text{red}}$ for all heights.

The fact is that, although $n_{0\,\text{green}}$ and $n_{0\,\text{red}}$ do differ by 3 in the sixth place, and although $n_{\text{green}}$ is greater than $n_{\text{red}}$ at all heights, it does not follow that the magnitudes of the two integrals are necessarily different. For illustration only, and not as a suggestion of fact, we may take the case in which $n_{\text{green}} = kn_{\text{red}}$, where $k$ is constant; it is obvious that in such an atmosphere $\psi_{\text{green}}$ would be exactly equal to $\psi_{\text{red}}$. That atmosphere would be achromatic. No green flash could be seen. There are countless ways in which $n_{\text{green}}$ and $n_{\text{red}}$ could be made to vary with height (even with the restriction that $n_{\text{green}}$ must be greater than $n_{\text{red}}$ for all heights) while still satisfying the condition of achromatism. One has only to look at the form of the integral to see that a change of only a few units in the ninth decimal place for the bottom layers of the atmosphere (say up to 200 feet) is sufficient to alter the value of the integral by 20″. In fact the differential changes between $n_{\text{green}}$ and $n_{\text{red}}$ that will change the condition of achromatism, where there is no green flash, into one favourable to the green flash, involve the remoter decimal places of the refractive index that are beyond the reach of direct measurement.

I suggest that whereas the author's explanation fails to account for the occasions when there is no green flash, the alternative outlined above does at least allow for all possibilities. If the normal condition is for the green flash to be an infrequent phenomenon, the conclusion must be drawn that the atmosphere is normally achromatic. The rarer occasions when the green flash is seen are departures from the normal. I am almost tempted to go further and suggest that the abnormalities conducive to the green flash are most probably found in the lower levels of the atmosphere. That is a conjecture that could be verified only if the refractive index could be measured with a much higher order of accuracy than is at present attainable. But in support, though not in proof, of my contention I would like to refer to measurements of the dip of the sea horizon made by ships of the Navy a few years ago. In this case abnormalities of the dip were entirely due to an abnormal condition of an extremely small part of the atmosphere, some seven or eight miles horizontally and 50 or 60 feet vertically. The normal condition of the atmosphere shows that a ray, nearly horizontal, has a curvature of about 8″ per mile. Our measurements of the dip, some three thousand in all, showed variations between 32″ and − 17″ per mile. If there are these wide variations in $dn/dh$, which is proportional to the curvature, is it not extremely likely that there may be differences between $n_{\text{green}}$ and $n_{\text{red}}$ in these lower levels sufficient to account for a green flash?

Mr W. H. White. I watch for the green flash in clear sunsets, in the mid-field of × 8 Zeiss binoculars, and of course avoid staring at the sun beforehand. On one exceptional September day, intensely clear from the previous sunset (green flash) in the Channel, the broken flakes of the setting sun in Essex showed yellow-green upper margins: all other observations have been at sea (where it would be more generally seen of men if sunset were not usually during dinner) though a small distant island was once no hindrance. Afternoon visibility must have been perfect, and the sun must begin to set with sharp outline, and without dulling to redness at the lower limb. A red sun, though it may spread at the last 3 or 4 sec. into a short line of liquid light, does not change colour. Change of colour begins at the horns of the thin segment, and moves inwards; it may be detectable 15 sec. in advance. A full orange sun goes pallid, and very briefly yellowish green; a yellow sun soon shows green edges and turns green to signal-green for 2 sec., or perhaps momentarily blue. With a bright, very slightly tinted, sun one may expect blue. In any case, the flash is many times brighter than the background of sky; and I have not noticed this ever producing complementary-colour illusions with white spars or the like. My last two observations were as follows:

(i) 1933/8/11, 33° N., 42° W. (400 miles S.W. Azores) ending a halcyon day of brilliance and calm, marvellously glassy blue sea. Green flash, turning blue, seen from foredeck (20 ft.) and exactly repeated from forecastle deck, 9 ft. higher, 10 sec. later; a longer interval than simple theory suggests. Sky not notably coloured (clear night, abundant shooting stars).

(ii) 1933/9/7, 26° 30′ N., 68° W. (300 miles S.W. Bermuda) moderate East wind, sea, and swell; western sky a broad blaze of bright gold, sun scarcely tinted. Green instantly turned blue, and lay as a liquid F line, persisting 5 sec. Not repeated on forecastle later. The air may have been slightly less warm than the sea.

Instructor-Lieut. D. K. McCleery. I should like to ask the author whether he can account for the fact that the phenomenon is much more frequently seen in the tropics than in higher latitudes. I have seen the flash in latitude 45° N. in the western Atlantic and it was a very brilliant one. There was on this occasion a double flash owing to the fact that I was standing in the bows and the ship was pitching to a W.S.W. swell. The temperatures of the air and sea were alike 56° F.; I mention this fact to show that there could not have been any abnormal refraction. It would be a good thing if observers could record the temperatures at sunset; they can easily be obtained from the ship's log.

The rate of setting of the sun is easily shown to be $15\sqrt{(\cos^2 \text{lat.} - \sin^2 \text{dec.})}$ seconds of arc per second of time, which means a rate of about 6 to 9 in our latitudes.

The conclusion at the end of the paper that "probably the green flash would not be seen with a blood-red sun" seems to be well borne out by experience. I have looked for the flash hundreds of times but have never seen it with a red sun.

Dr W. D. Wright. I have made a large number of experiments recently on the effect of successive contrast in changing the appearance of various colours and in no case has it been possible to change a blue into a green, provided that the blue has been reasonably pure and not a blue-green. The correct explanation of this is, apparently, that blue is one of the fundamental responses according to the trichromatic theory of colour vision, so that even if a deep violet radiation were available to produce the contrast, blue could still not be converted into green. Moreover the effects of successive contrast are very much stronger than the similar effects of simultaneous contrast, so that I do not think there is any doubt that the cause of the flash appearing green instead of blue is physical rather than physiological.

Mr A. F. DUFTON. In his introductory remarks the author referred to the green flash as a somewhat rare phenomenon. There is another phenomenon which is proverbially rarer still, the blue moon. Until this afternoon I did not realize that the moon could appear blue, but an hour ago I had occasion to refer to *Nature* for 1883 and found there an account of the green sun which in September 1883 was seen at places as wide apart as Madras, Trinidad and Ceylon. The account from Ceylon is interesting: "The sun for the last four days rises in splendid green when visible, i.e. about 10° from the horizon. As he advances he assumes a beautiful blue, and as he comes from this on looks a brilliant blue resembling burning sulphur.... Even at the very zenith the light is blue.... As he declines the sun assumes the same changes but vice versa...." The account goes on to state that the moon behaved in similar manner. The phenomena were attributed to the eruption in Java.

Mr J. GUILD. The evidence given in the paper leaves no doubt in the minds of physicists that scattering is adequate to account for the green colour of the flash. Prof. Haldane, whose views are quoted in the paper, appears to over-stress the importance of simultaneous contrast in determining the colour quality of visual sensations, and to leave out of account many other important factors. As a matter of fact it is obvious on quite general principles that the change from spectral blue to green cannot be due to simultaneous contrast. The contrasting colour would have to be of still shorter dominant wave-length and of extremely high intensity. No conditions at all resembling this exist in the present problem.

Mr T. C. SUTTON. My remarks concern merely the ease with which the green flash may be seen. Under suitable conditions, this can be done more frequently than some of the speakers suppose. Some years ago I was in latitude 40 S.; the house faced due west, 200 ft. above an ancient lake bed, now a fertile but dusty plain. A phenomenon which I believe to be this flash was seen with perhaps 60 per cent of the sunsets. It was not always green; it might be any shade between orange and blue, but not red. The atmosphere was usually dust-laden. After rain the sky was very blue, and the effect was then seldom, I think never, observed. During one visit to this house, soon after or at the time of the eruption of Ngauruhoe 2000 miles due east, many strange sights appeared, among them the blue moon which has been referred to. It is perhaps worth notice that, while dusty summer atmospheres tended to make the moon appear orange, the higher and finer volcanic dust turned it a pale but brilliant blue-green.

Mr W. F. FLOYD. I should like to ask whether the spectral distribution of the light from the lamp used in the experiments described in § 3 of the paper approximates closely to that of sunlight at sunset on occasions which seem to be favourable to the appearance of the green flash; and further what is the order of the ratio of the absolute intensities of the light reaching the observer's eye in the experimental and in the natural phenomena. I am willing to accept a purely physical explanation of the green flash, but I would suggest that the present paper establishes only one possible explanation. That a physiological effect may contribute to the phenomenon, or may even provide the sole explanation of it, has not been disproved, nor can it be disproved by purely physical arguments. As a contribution to the evidence that has been offered in the discussion, I would submit the following account of a personal experience. I have often looked towards the early morning sun from my window. If, having done so, I look on to a neutral background, or a coloured but not well illuminated background, I see an image of the window, the window space being full green in colour and the surrounding ground dark. I am able to see a similar image by closing my eyes and shielding them from direct light from the window, but in this case more detail is discernible, the image of the window space being fringed with blue-violet or violet. The succeeding events, while not bearing directly on the present discussion, are of interest. I find that with the eyes closed I can hold the image for several minutes by controlling the intensity of the light falling on the closed eyes, and also that the image colours change slowly down the spectrum. Later in the day I have found that the full green stage was preceded by a pale blue stage.

AUTHOR's reply. So far as I understand Captain Baker's point of view, he thinks that the calculation of atmospheric dispersion given cannot be correct, because it proves too much; if it were correct, the green flash would always be seen, whereas it is in fact only rarely seen in most places. Most previous writers have felt more difficulty in satisfying themselves that the dispersion is enough. Captain Baker gives reasons which lead him to think that the atmospheric refraction may be of an achromatic character, but I do not clearly understand how achromatism can be got out of a single refracting medium: unless indeed variations of composition of the air are relied on, and I doubt if they can be stretched far enough.

But the whole difficulty seems to me illusory. False light in one form or another is an amply sufficient explanation of the cases in which green flash is not observed. In my experiments with the water prism I found it most necessary to take every precaution against false light, and Nature takes no such precaution.

In reply to Mr McCleery: I have always supposed that the comparatively common occurrence of the green flash in the tropics is to be attributed to the clearness of the horizon at sunset. The same would apply particularly to a desert horizon, and judging from the experience of one or two archaeologists who have lived in the desert for long periods, the green flash is a common occurrence there.

In reply to Mr Dufton: The blue sun seen under certain conditions is of great interest, but I am not able to throw any light upon it.

In reply to Mr Floyd: The lamp used was of the same order of intensity as the setting sun; but, in the absence of a scattering medium, it was bluer than the setting sun usually is. With the scattering medium, the colour passed through the various shades observed in the setting sun on different occasions.

535·243

# AN ULTRA-VIOLET PHOTOELECTRIC SPECTROPHOTOMETER

By D. H. FOLLETT, M.A., A.Inst.P., Adam Hilger, Ltd.

*Communicated by F. Twyman, F.R.S., December* 22, 1933. *Read April* 20, 1934.

*ABSTRACT.* An instrument designed for accurate spectrophotometry is described. The transmission of the sample is found by comparison with a rotating sector, which is of cylindrical type and meets the criticisms usually directed against sectors whose transmission can be varied while they are running. Two photocells are employed and fluctuations in the intensity of the source are compensated for. The possible sources of error and their elimination are discussed and the results of tests are given.

## § 1. GENERAL CONSIDERATIONS

In certain applications of spectrophotometry the use of photoelectric methods is definitely preferable to the use of photographic methods⁽¹⁾. No great confidence is felt, however, in photoelectric methods in general, as their absolute accuracy is often in question although they may be highly sensitive. In the instrument here described special attention has been given to the question of accuracy as distinct from sensitivity, and we consider that all the predictable causes of error have been avoided. Nevertheless, we are well aware that theoretical considerations are of limited value in appraising the worth of any photoelectric arrangement and that the only true test is that of comparison with other well-tried methods.

The use of gas-filled photoelectric cells is more or less essential owing to the necessity for obtaining a high sensitivity. This immediately precludes the use of any method in which the photocell is used actually to measure the intensity on the assumption of a linear relation between current and intensity. Spectrophotometric methods have been described in which this relation is assumed for gas-filled cells, and for which the authors claim that accurate results are yielded. They may have been fortunate in obtaining such cells, however, and it seems better to avoid such an assumption if possible.

As ultra-violet sources are most conveniently run from the mains supply, it is necessary to use some method in which fluctuations in intensity can be compensated for; this would probably be necessary even if the electric supply were perfectly steady. Smith and Holiday(2) have described how they stabilized the output from a hydrogen tube, but as perfect stabilization cannot be obtained it seems desirable, since the highest accuracy is aimed at, to use a method which compensates for fluctuations even when a stabilizing device is used.

The method finally adopted is a modification of one which has been described by von Halban and Ebert(3) and has been shown by von Halban and Eisenbrand(4) to give results agreeing closely with those given by various photographic methods. The fact that their arrangement gave accurate results does not, of course, provide any guarantee that the present instrument will do likewise, but it indicates that there is no fundamental reason to the contrary.

## § 2. MEANS OF VARYING LIGHT INTENSITY

It is necessary to have a means of varying light intensity in a known way, and we have used a rotating sector for this purpose. The validity of Talbot's law for photoelectric cells has been questioned by some workers, but opinion seems fairly unanimous in its favour in England(5). A more serious difficulty in the use of sectors is that they should be adjustable when actually running, and that owing to wear in the rather complicated mechanical arrangements by which this is achieved loss of calibration soon occurs. We have avoided this difficulty by using a sector of a new type described by Dunn(6). This consists of a cylinder, rotatable about its axis, with two apertures cut in its surface. These apertures are similar, and are bounded on one side by a straight spiral cut round the surface, and on the other by an edge parallel to the axis of the cylinder. Figure 1 (*a*) shows the sector, and figure 1 (*b*) shows the surface developed to a plane. *ab*, *a′b′* are the spiral edges and *ac*, *a′c′* the edges parallel to the axis. A beam of light passing along the axis of the cylinder is reflected radially by a right-angle prism. The cylinder is movable along its axis, and as it rotates, the light alternately passes through the aperture and is obstructed by the wall. The fraction of a revolution during which the light is transmitted depends upon the position of the sector with respect to the deflected beam. When the cylinder is moved along its axis the change in the fraction transmitted obviously varies linearly with the amount of movement.

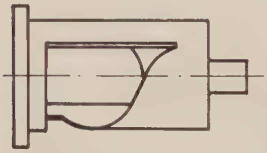


Figure 1 (*a*).

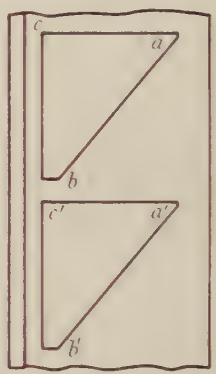


Figure 1 (*b*).

## § 3. PRINCIPLE OF THE METHOD

Figure 2 indicates diagrammatically the optical system of the photoelectric arrangements. The lens *B* collimates the beam after it emerges from the slit *S* of the monochromator. The quartz plate *H* reflects a fraction of the beam into a photocell

$L$, and the rest of the beam passes through the absorption tube $C$ containing the sample. Lens $K$ then focuses the beam, after reflection by the prism $P$, so that an image of the slit is formed on the wall of the sector. The light then passes into the photocell $F$. The photocells are connected so that they deliver currents of opposite sign to a Lindemann electrometer. The potentials on the cells are so adjusted that, with the sector running in its fully open position, these currents are exactly equal so that the electrometer does not charge up. The absorption cell $C$ is then slid out of the beam and replaced by a comparison cell $C_1$ (e.g. one containing the solvent). The circuit is then re-balanced by adjusting the position of the sector. The photocell $F$ is now receiving the same amount of light as before, and the ratio of the transmissions of the sector in the two settings gives the ratio of the transmissions of the two absorption cells $C$ and $C_1$. No relationship between current and intensity is assumed. Since both photocells are illuminated by monochromatic light any variation in the intensity of the source will not affect the balance provided that both cells have the same (intensity, current) characteristic. This condition will be closely fulfilled over the small range of intensities in which fluctuations are likely to occur.

## § 4. OTHER METHODS CONSIDERED

Since the production of this instrument a description has been published of an arrangement which might be equally satisfactory(7). A vacuum cell is used, and the linear relation between current and intensity is assumed. As no sectors are used, the new electrometer valve is employed to indicate null points. This is probably capable of higher sensitivity than the Lindemann electrometer, but the method entails the use of a high-resistance potentiometer ($10^8\,\Omega$ in the arrangement described) and the sensitivity depends upon the value of this resistance. There are obvious difficulties in the manufacture of a high-resistance potentiometer, and it is probable that $10^8\,\Omega$ is about the highest that can be used. Under these conditions the sensitivity is little, if any, greater than that of the Lindemann electrometer. The use of the gas-filled cell, possible in our method, thus seems to give a definitely higher sensitivity.

Another method of compensating for fluctuations in the light-source is that described variously by Hardy(8), Dobson and Perfect(9) and others. The beam emerging from the monochromator passes alternately through the sample and through a device for varying the intensity, the frequency of the alternation being from 10 to 50 times a second. The light falls on a photocell so coupled to an amplifier that only fluctuating currents are transmitted. If the two beams are not equal the light falling on the cell will be flickering, so that the amplifier gives an output, and the variable beam is adjusted until the amplifier shows no output. This method sounds very simple and large fluctuations of intensity in the original beam can be dealt with, provided that their period is large compared with that of the alternations. For ultra-violet work the chief difficulty in applying the method lies in finding a suitable means of varying the intensity of the comparison beam in a known manner; the rotating sector in some form, or a variable aperture, are the only two feasible means in general, and of these the former is inapplicable to this particular case. The latter is always difficult to apply in photoelectric apparatus because an image of the monochromator slit has to be formed on the surface of the cell; quite slight aberration will cause variation in the size of this fine image for different apertures, and unless the sensitivity of the cell surface is completely uniform error is introduced in this manner.

## § 5. INSTRUMENTAL DETAILS

The lenses $B$ and $K$, figure 2, are of quartz and it is therefore necessary that a focusing movement should be provided; knobs working over engraved scales are supplied for this purpose. The absorption tubes are mounted on a slide and are pushed alternately into the beam by means of a further knob. The photocells are Osram UNG 7, which is gas-filled and has a sodium cathode and quartz windows. The electrometer is read by means of a microscope and is illuminated by a lamp mounted outside the case; it is of the Lindemann type made by the Cambridge Instrument Company, Ltd., and is chosen for its robustness and quick period. The earthing-switch is built into the electrometer and is operated pneumatically by means of a rubber bulb. The resistances used are ordinary commercial wireless components; we have used these for two or three years now and have never encountered any trouble which could be traced to them.

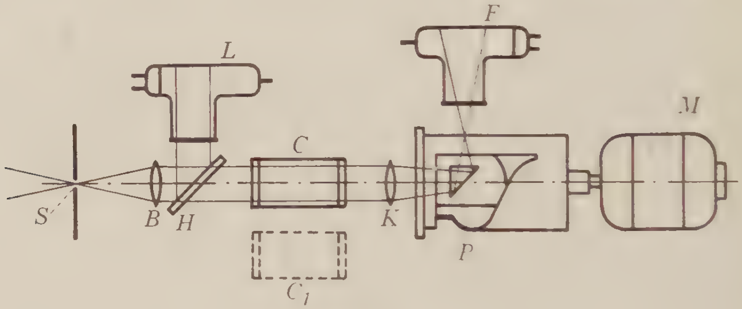


Figure 2.

The sector itself is a cylinder 3 in. in diameter and the spirals are of 6-in. pitch. The closed end $a$ does not go completely to a point, but has an aperture of about $\frac{1}{30}$ of the value of that at the open end $c$. It is mounted directly on the shaft of a motor $M$, which was specially selected to have no end play and to run without vibration. The only suitable motor which we could find is an Italian one, made by the firm of Marelli. It is mounted on a carriage which slides along the optical axis of the system, and is moved by a heavy screw of $1\frac{1}{2}$ mm. pitch, which enables the movement to be measured. The number of complete turns of the screw is registered by a revolution-counter, and the additional fraction of a turn is read by means of the graduations on the edge of a drum.

The photoelectric circuit is shown in figure 3. The resistances $R_1$, $R_2$, $G_1$, $G_2$ are the usual protective resistances for the photocells and the electrometer respectively. $A_1$ and $A_2$ are the potentiometers which vary the potentials on the cells; $B_1$ and $B_2$ are the potentiometers which vary the potentials on the plates of the electrometer. The circuit differs slightly from the orthodox circuit in that no connexion is made from the point $E$ to earth. The current through the potentiometers $B_1$ and $B_2$ is then the same, and any variation in the batteries does not affect the ratio of the potentials on the two plates. This is the condition that the electrical zero shall remain constant⁽¹⁰⁾. If $E$ is connected to earth, a change in the battery $F$, for example, will alter the potential on one plate only, so that the electrical zero will alter. It is important in this instrument that no changes of zero should occur while the two settings constituting one reading are being made. A similar arrangement is made for the potentiometers controlling the potentials on the two photocells.

The change $\delta A$ in aperture produced when the sector is moved a distance $d$ is given by $\delta A = kd$, where $k$ depends on the pitch of the spiral. The distance $d$ can be found from the indication of the revolution-counter and the drum. However, the actual aperture of the sector must be known, since the ratio of the apertures at which balance is obtained for both sample cell and comparison cell must be found. The actual aperture is given by the relation $A = kd + P$, if $d$ is the reading of the position of the drum and $P$ is a constant which is the reading for the position at which the aperture would be zero if the sector closed down to a point at $a$, figure 1 ($b$). The value of $k$ is easily calculated from the pitch of the spiral. $P$ is most easily found as follows. A measurement is taken of the absorption of any convenient sample at any wave-length. Suppose $A_1$ and $A_2$ are the apertures of the sector at the two balance points. Then another reading is taken, this time with the apertures approximately half what they were before. If $d_1$, $d_2$ and $d_1'$, $d_2'$ are the respective readings of the position of the sector, then

$A, d$
$k$

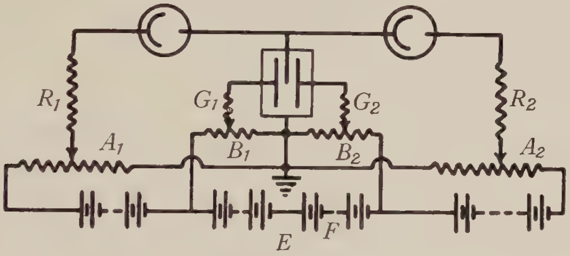


Figure 3.

$P$
$A_1, A_2$
$d_1, d_2,$
$d_1',$

$$\frac{A_1}{A_2} = \frac{kd_1 + P}{kd_2 + P} = \frac{kd_1' + P}{kd_2' + P}.$$

$k$ being known, $P$ can then be calculated.

## § 6. POSSIBLE SOURCES OF ERROR

The most obvious source of error was in the determination of the constants of the sector. If the constant $P$ is zero, the value of $k$ is immaterial. $P$ can be made very small by suitably setting the revolution-counter, and the value of $k$

can be calculated with sufficient accuracy from the pitch of the spiral. In this instrument we thought it advisable to measure the aperture independently, and this measurement was carried out in the following way.

A beam of light from a motor head lamp was passed through the sector, and then through a radial sector and into a photocell connected to a galvanometer. The cylindrical sector was held stationary so that light passed through it uninterrupted, and the radial sector was rotated. The deflection of the galvanometer was noted, and then the sector was stopped. The radial sector was then set so that light passed through it uninterrupted; the cylindrical sector was set rotating and its position was adjusted until the galvanometer gave the same deflection as before. The angular aperture of the radial sector was measured carefully to about 5''. The position of the cylindrical sector was plotted against the aperture of the radial sector, and the gradient of the plotted line was determined. It was found to be 0·125 radian per turn of the screw, whereas calculation from the ostensible pitch of 6 in. gives 0·124 radian per turn. This difference of nearly 1 per cent may arise in the setting up of the machine for cutting the spiral, or owing to the straight lines *ac*, *a'c'*, figure 1, not being cut exactly parallel to the axis of the cylinder.

Further experiments were carried out to test the uniformity of the curved edges of the sector. A mirror was fixed to the surface of the sector with wax, and a beam of light was reflected from it to a scale at a distance of 1 m. A microscope was then focused on the spiral edge and the position of the sector was set so that the edge came on to the cross webs. The sector was traversed a short distance, and then rotated until the edge came again on to the cross webs in the microscope. The movement of the beam on the scale was measured, a number of readings were taken and the scale readings plotted against the movement of the sector. Only a small section of the sector could be dealt with in one run of readings, otherwise the deflection became too big. Small local irregularities were found, but were completely negligible except perhaps at the narrow end of the sector. The magnitude of these random errors is never more than 0·0019 radian, and the smallest possible angular aperture is 0·200 radian for one aperture, corresponding to a transmission of nearly 7 per cent. Thus, even if one of these random errors occurs in the same direction in each aperture at any one setting of the drum, the error cannot be more than 1 per cent. For higher settings the error is proportionately less.

The error due to inaccuracy in the value of $P$, in the expression for the aperture, is not reduced by making $P$ small, since $P$ can be determined only to within certain limits which are independent of its actual value. Nevertheless, it is an advantage, as has been stated above, to make $P$ small in order to increase the tolerance in $k$.

Error in positioning is avoided by joining the photometer to the monochromator by means of a ball-and-socket arrangement in virtue of which the photometer can rotate slightly, in a horizontal plane, about the exit slit of the monochromator when the two instruments are joined up. The exact adjustment is not critical and is made very simply as follows.

Figure 4 represents the optical system. An image $I$ of the slit $S$ of the monochromator is formed on the surface of the sector. The only possible displacement of the photometer is a slight rotation about $S$, and the effect of such rotation is to cause the radiation leaving the slit to enter the photometer obliquely. The dotted lines represent rays so entering. If aberrations are absent, the image $I$ is, of course, formed in the same place on the sector regardless of the angle these rays make with the axis, the only effect being a loss of light. The lining up is therefore not critical and is obtained very simply by slipping a tube over the end of the mount of the lens $K$. On the end of the tube are engraved a pair of parallel lines equidistant from the axis, and the instrument is rotated until the beam falls centrally between them.

The lenses $B$ and $K$ have a focusing motion along the axis. It is very desirable that they should be so mounted that the motion is accurately along the axis. If it is not so, there will be a displacement of the image $I$ as the lens moves. Thus the value of the constant $P$ in the expression for the aperture will be different for different wave-lengths. This is undesirable, although the value of $P$ could be found for all wave-lengths, since it would entail additional calibration. The presence of the defect was looked for by focusing a microscope on the image $I$, and observing whether this moved as the wave-length of the monochromator radiation was changed and the lenses $D$ and $C$ were focused.

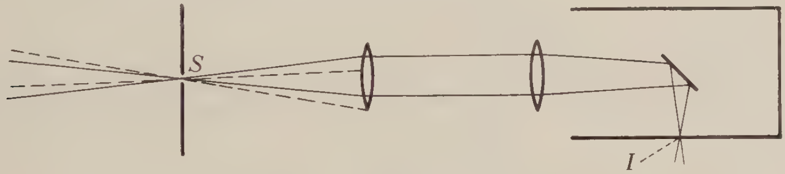


Figure 4.

Another possible source of slight error exists, though it is rather obscure. The width of the image on the sector is finite, so that the aperture of the sector varies across the slit. Let figure 5 represent the image of the slit on the sector. Let the line *ef* be such that the effective aperture of the sector is the same along the whole length of it, and let *gh* be a line parallel to it and distant $s$ from it. Let the intensity be uniform along it and equal to $I$. Then it can easily be shown that the line *fe* represents the effective position of the whole image on the sector if $\Sigma Is = 0$. Obviously for a rectangular image uniformly illuminated, the effective position is the line dividing it into two geometrical halves. In practice, however, the image will not be so simple as this. The monochromator used with the photometer was the Müller-Hilger double monochromator. The front slit of this is curved, the curvature being such that a line of wave-length 3000 Å. gives a straight image at the slit which divides the two monochromating systems. This slit is straight. For any other wave-length the image at the slit will be slightly curved. To illustrate the point we may refer to figure 6. $j$ represents the shape of the first slit. The full lines in $k$ represent the shape of the image after dispersion in the first system. The dotted lines represent the boundaries of the slit, which cuts off a small portion of the image, as can be seen so that the shape becomes that shown at $l$. After dispersion by the second system the shape becomes as shown at $m$. The left-hand edge will fit the curved exit slit, but the right-hand will not. The difference between the shape of the slit and that of the image will be different for different wave-lengths, so that in effect the value of the constant $P$ will depend upon the wave-length. The extent of the effect will depend upon the length of the slit and can be reduced to negligible proportions by confining the illumination to the centre of the slit. With a continuous spectrum the effect practically disappears, because the space between the full and dotted lines (representing the left-hand edge of the slit) in $m$ will be filled in. Thus the whole width of the slit will be illuminated, and the departure from uniformity of illumination will be very small except in exceptional cases where the absorption band is so steep that there is a marked change in absorption for the wave-length within the range which the slit covers. In this case there will be error due to the finite width of the monochromator slits in any case.

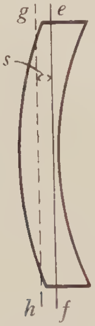


Figure 5.

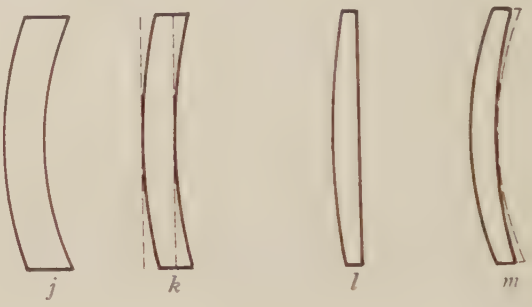


Figure 6.

## § 7. TESTS

To see if there were any systematic errors inherent in the instrument the transmissions of two pieces of optically worked neutral glass were separately measured, and then the two glasses were put together and the transmission of the pair was measured. The results are given below. Each measurement is the mean of several readings. Each figure is given to 4 places, though the significance of the fourth figure is very doubtful.

Transmission of first sample ................0·2749.
Transmission of second sample ..............0·2746.
Calculated transmission of the two together ....0·07548.
Observed transmission of the two together .....0·07558.

The difference between the observed and calculated values is 0·13 per cent of their value.

The pieces of glass here used had been measured by the National Physical Laboratory at wave-length 0·5461$\mu$. We have measured one piece at this wave-length and find an excellent agreement. The National Physical Laboratory figure for the transmission is 0·2985, while our instrument gives 0·2981.

## § 8. ACKNOWLEDGMENTS

Acknowledgments are due to Messrs Adam Hilger Ltd., at whose suggestion the paper is published; to Mr F. Twyman, F.R.S., and Dr F. Simeon who have contributed advice and suggestions; and to my assistants Mr W. J. McCarthy and Mr R. A. C. Isbell who have performed most of the experimental work in the development of the instrument.

## REFERENCES

(1) For a discussion of the relative applicability of photographic and photoelectric methods see (4) below, and TWYMAN, *Proc. Phys. Soc.* **45**, 643 (1933).
(2) SMITH and HOLIDAY. *Trans. Opt. Soc.* **33**, 9 (1931).
(3) VON HALBAN and EBERT. *Z. f. Phys. Chem.* **112**, 321 (1924).
(4) VON HALBAN and EISENBRAND. *Proc. R.S.* **116**, 153 (1927).
(5) CARRUTHERS and HARRISON. *Phil. Mag.* **7**, 792 (1929); J. Guild in the discussion on paper by Campbell Smith and Holiday, (2) above.
(6) DUNN. *Rev. Sci. Inst.* **2**, 807 (1931).
(7) WINN. *Trans. Faraday Soc.* **29**, 689 (1933).
(8) HARDY. *J. Opt. Soc. Amer.* **18**, 96 (1929).
(9) DOBSON and PERFECT. *Phys. and Opt. Socs., Discussion on Photoelectric Cells*, p. 79 (1930).
(10) GRIMMETT. *Prcc. Phys. Soc.* **45**, 117 (1933).

## DISCUSSION

Mr J. GUILD. Much attention has been devoted to the design of photoelectric spectrophotometers in recent years, and extravagant claims have been made for their performance which, on examination, have usually been found to be quite baseless. The sources of possible systematic error in such apparatus are often very difficult to spot on theoretical grounds, and for this reason I never, nowadays, judge of the merits of an instrument of this kind from consideration of the design, but only from the results of critical comparisons with other instruments of known reliability. I could have wished, therefore, that the author had included more experimental results. However, in giving two he has given two more than is usual, and the results of these two certainly suggest that the present apparatus is not only sensitive but, what is much more important, accurate.

Mr T. SMITH. It is of interest to examine the test mentioned by the author in which the transmissions of filters together and independently are compared. The filters should of course be definitely separated from one another, and we may assume that they are placed normal to a parallel beam of light. The mutual effects of a single filter and the instrument should differ very little from those of two filters, and we can, therefore, consider two filters alone as an absorbing system.

Assume that reflection takes place only at the surfaces of the filters, and that in the body of the filter only absorption occurs. Then if $P$ and $N$ denote the amounts *P, N*

of light travelling in the positive and negative directions on the incident side of one
$P', N'$ filter, and $P'$ and $N'$ are the corresponding quantities on the emergent side, we have

$$(PN) = (P'N')\, m,$$

$m$ where

$$m = \begin{pmatrix} 1+\rho & \rho \\ -\rho & 1-\rho \end{pmatrix} \begin{pmatrix} \frac{1}{\tau} & \cdot \\ \cdot & \tau \end{pmatrix} \begin{pmatrix} 1+\rho & \rho \\ -\rho & 1-\rho \end{pmatrix},$$

the three component factors corresponding to surface, body and surface of the
$\rho$ filter respectively; $\rho$ is the ratio of the reflected to the transmitted light for one
$\tau$ surface only, and $\tau$ is the transmission factor for the body of the filter. On putting
$N' = 0$ we get as the total transmission factor measured by a correct method

$T$
$$\frac{P'}{P} = \left\{\frac{(1+\rho)^2}{\tau} - \rho^2\tau\right\}^{-1} = T \text{ say}.$$

$m_1$ If now this filter is followed by a second for which the matrix is $m_1$, the fundamental equation is

$$(PN) = (P''N'')\, m_1 m,$$

and the combined transmission on the assumption that the second filter only differs from the first in transmitting $\tau_1$ instead of $\tau$ is

$$\frac{P''}{P} = \left[\left\{\frac{(1+\rho)^2}{\tau} - \rho^2\tau\right\}\left\{\frac{(1+\rho)^2}{\tau_1} - \rho^2\tau_1\right\} - \rho^2\left\{\frac{1-\rho}{\tau} - (1-\rho)\,\tau\right\}\left\{\frac{1+\rho}{\tau_1} + (1-\rho)\,\tau_1\right\}\right]^{-1}$$

or

$$TT_1\left[1 - \left(\frac{\rho}{1+\rho}\right)^2 (1+\tau T)(1+\tau_1 T_1)\right]^{-1}.$$

The quantities measured are $T$ and $T_1$. We can assume a value for $\rho$ with sufficient accuracy, and for $\tau$ substitute the approximate value $T(1+\rho)^2$. The expression for the computed joint transmission is then

$$TT_1\left[1 - \left(\frac{\rho}{1+\rho}\right)^2 - \rho^2(T^2 + T_1^2) - \rho^2(1+\rho)^2\, T^2T_1^2\right]^{-1}.$$

The author's calculated transmission 0·07548 has been obtained on the assumption that no reflected light enters into the problem. If we apply the theoretical corrections for refractive index $\mu$ we have $4\mu\rho = (\mu - 1)^2$, and this gives the following calculated transmissions, when

$$\mu = 1{\cdot}5, \quad \text{transmission} = 0{\cdot}07563,$$

and when

$$\mu = 1{\cdot}6, \quad \text{transmission} = 0{\cdot}07574.$$

Either of these values changes the sign of the difference between the observed and calculated results. The refractive index of the specimens used in this test is likely to be nearer 1·5 than 1·6, and as the surfaces were not freshly polished at the time of the test the reflected light would be somewhat less than the value derived from the refractive index. A transmission value close to 0·07560 is as likely as any we might suggest, and the comparison of this figure with the observed value 0·07558 leaves nothing to be desired.

Dr W. D. Wright. It would be interesting to know whether Mr Follett's experience with the cylindrical sector he describes would lead him to recommend it for general use in the laboratory. The calibration evidently depends on the position of the various optical parts in relation to this sector; thus, any transverse movement of the light beam or longitudinal change of focus would upset the calibration. While it is not difficult to ensure that such movement shall be absent in a complete instrument made in a factory, the same does not always follow in apparatus assembled in the laboratory.

Author's reply. I entirely agree with Mr Guild's view that the merits of a photoelectric spectrophotometer should be judged only from the results of comparisons with instruments of known reliability. This statement may seem inconsistent with the fact that only two results are given, but unfortunately we had no time before delivering the instrument to carry out all the tests we should have liked to make, and the results of the two tests given in conjunction with the test on the sector itself, as detailed in the paper, seemed so satisfactory that we felt justified in letting it go. We are hoping to put together another exactly similar instrument shortly; further tests will then be applied, and the results will be published.

I am grateful to Mr Smith for the details of the correction to be applied to the calculated values of the transmission of the two glasses together. Actually I had not overlooked the possible necessity of a correction, but a rough calculation satisfied me that the effect was negligible. Evidently my calculation was rougher than I thought, but I am relieved to find that the modification does not spoil the agreement between the calculated and observed values. Actually there is no experimental justification for the fifth figure in the results I gave—it would have been better to express the figure to four decimal places only; as it stands it may give an exaggerated impression of the sensitivity of the instrument. This does not affect the value of Mr Smith's remarks, of course; but the present seems a suitable opportunity of calling attention to the point.

I agree with Dr Wright that absolute rigidity of the mounting of the sector and optical system is necessary if accurate results are to be obtained with the cylindrical sector. It follows that the sector is hardly suitable as a general-purposes accessory for laboratory use; its proper use is in a permanent assembly.

532.612.4:662.232.1

# NOTE ON THE PURITY REQUIRED FOR SURFACE-TENSION MEASUREMENTS

By T. C. SUTTON, M.Sc., F.Inst.P., and H. L. HARDEN, B.Sc.,
Research Department, Royal Arsenal, Woolwich.

*Received January* 25. *Read April* 20, 1934.

*ABSTRACT.* The extent to which impurities affect the measured value of the surface tension of a liquid depends on the method of measurement employed. This effect is applied to test whether the purity of a sample is adequate for the measurement of the surface tension of the pure liquid.

In the course of measurements of the surface tension of some liquid explosives, the well-known differences between the values obtained for the surface tension by dynamic and by static methods of measurement and the effect of impurities on these values⁽¹⁾ have been applied as a means of checking whether the purity of a sample is adequate for measuring the surface tension of the pure liquid.

The table refers to measurements made during the purification of an organic liquid (nitroglycerine), analysis of which was both difficult and not sufficiently sensitive. The effect of residual impurities is conspicuously greater with the capillary methods than with the bubble-pressure method. Agreement between the three methods is an indication that (i) the purity is adequate and (ii) the value obtained is that of the pure material.

| Surface tension (dyne/cm.) | | | Remarks |
|---|---|---|---|
| Bubble-pressure method | Capillary methods | | |
| Sugden's(2) | Ferguson and Kennedy's(3) | Sutton's(4) | |
| 49·1 | 43 | 38 | After preliminary drying and filtering |
| 49·8 | 45 | 42 | After further drying and filtering |
| 50·3 | 50 | 51 | After repeated drying and filtering. Value unaltered by further treatment |

A sample of 2·0 cm³ was sufficient for test by these three methods. The test may be applied, however, to much smaller quantities of material by suitably selecting two capillary methods for the purpose.

This test of purity does not consume the sample.

## REFERENCES

(1) P. L. de Noüy, *Phil. Mag.* **48**, 264 (1924); F. Schmidt and H. Steyer, *Ann. der Phys.* **79**, 442, 5 (1926); O. K. Rice, *J. Phys. Chem.* **31**, 207 (1927); and others.
(2) S. Sugden, *J. Chem. Soc.* **125**, 27 (1924).
(3) A. Ferguson and S. J. Kennedy, *Proc. Phys. Soc.* **44**, 511 (1932).
(4) T. C. Sutton, *Proc. Phys. Soc.* **45**, 88 (1933).

## DISCUSSION

Dr A. Ferguson. Obviously it is not necessary to employ the three methods to test the purity of the material under consideration—the three methods are approximately of the same order of sensitiveness, and any one of them is adequate. It is only necessary to purify the material under test until a constant value is obtained.

The figures given open up a number of important considerations, especially that of the variation of surface tension with time. This question has not received the consideration which it merits, and one of my students has recently made a study of the relation between surface tension, concentration, temperature and time. The solutions used were colloidal in character and a number of hitherto unsuspected relations have been brought to light.

Reply by Mr T. C. Sutton. I could agree entirely with the remarks of Dr Ferguson were purification always a simple process and were its results always determinate.

In many cases, however, such as those with which we have been dealing at Woolwich, there may be no ready means of telling when a process of purification has yielded a pure sample and when, for example, a "constant-composition mixture" has been obtained. Such mixtures—for instance, certain mixtures of homologues—have frequently been found to be beyond the powers of analysis, and may easily be mistaken in practice for pure substances. Some further test was desired, and we were glad to find that the relevant purity of our sample could be checked by an expeditious means independent of the method of purification employed.

621.397.5

# VELOCITY MODULATION IN TELEVISION

By E. E. WRIGHT, Baird Television, Ltd.

*Communicated by Dr Lewis Simons, January 17, 1934. Read April 20, 1934.*

*ABSTRACT:* The motion of a spot of light of constant intensity, necessary to produce the effect of a sinusoidal linear distribution of light-intensity on a television viewing screen, is determined and the effect of the finite size of the scanning spot is compared with the analogous effect in the more usual type of television system in which the scanning spot moves with constant speed and is modulated in intensity.

## § 1. INTRODUCTION

According to the method of velocity modulation, a desired visual or photographic effect of light and shade-distribution is obtained by moving a spot of light of constant intensity along a scanning-line at a varying velocity.

If the size of the light-spot is small compared with the details to be reproduced, the motion required to produce a given shade-distribution is approximately obtained by making the velocity at any point inversely proportional to the required brightness at that point. If, however, the size of the light-spot is comparable with the details, then a more refined analysis is required to determine the necessary motion.

## § 2. THEORY

$W$ If a rectangular light-spot of unit intensity and width $2W$ moves so that its
$x, t$ median line is displaced through a distance $x$ at a time $t$ and if the velocity is always
$f$ positive, then the motion necessary to give a shade-distribution $f(x)$ is defined by the equation

$$\int_{x-W}^{x+W} \frac{dt}{dx} . dx = f(x) \qquad \text{......(1).}$$

In this preliminary analysis a simple periodic distribution defined by

$A, B$

$$f(x) = A + B . \cos x$$

is considered. With this distribution equation (1) will be satisfied when

$$t = Ax/2W + \tfrac{1}{2}B . \text{cosec}\, W . \sin x \qquad \text{......(2),}$$

and in a sense this solves the problem of determining the motion, but in order to draw conclusions it is necessary to obtain $x$ explicitly in terms of $t$.

$T, C$ Putting $T = 2Wt/A$ and $C = (WB . \text{cosec}\, W)/A$, we reduce equation (2) to

$$T = x + C . \sin x \qquad \text{......(3).}$$

If $x$ is increased by $2\pi$ then $T$ is also increased by $2\pi$, and if the sign of $x$ is reversed, so is the sign of $T$, hence if

$$x = T + \phi(T) \qquad \text{......(4)},$$

it follows that $\phi(T)$ is a periodic odd function, of period $2\pi$, and is therefore adequately given by

$$\phi(T) = \sum_{n=1}^{\infty} A_n \sin nT \qquad \text{......(5)}.$$

Now, $$\pi A_n = \int_0^{2\pi} x.\sin nT.dT - \int_0^{2\pi} T.\sin nT.dT = \frac{1}{n}\int_0^{2\pi} \cos nT.dx \qquad \text{......(6)},$$

by integration by parts and use of equation (3).

The integral in equation (6) is easily reduced to a known form* and we have

$$\begin{aligned} \pi n A_n &= \int_0^{\pi} \cos(nx + nC.\sin x)\,dx + (-1)^n \int_0^{\pi} \cos(nx - nC.\sin x)\,dx \\ &= \pi J_n(-nC) + (-1)^n \pi J_n(nC) = 2(-1)^n \pi J_n(nC) \qquad \text{......(7)}, \end{aligned}$$

where $J_n(nC)$ is a Bessel coefficient of order $n$. Hence the equation of the necessary motion is

$$x = T + \sum_{n=1}^{\infty} \frac{2(-1)^n}{n} J_n(nC).\sin nT \qquad \text{......(8)}.$$

In practice the constant-velocity part of this motion is produced by a local oscillator generating saw-toothed wave-form at the receiver, and a signal to produce the periodic part only is transmitted.

## § 3. COMPARISON OF FINITE-APERTURE EFFECTS

In an ordinary television system, in which the scanning spot of width $2W$ moves with a constant velocity $V$ and is modulated in brightness, it can be shown that†, owing to the finite size of the aperture, in order to produce a periodic distribution of shade $B.\cos x$ it is necessary to transmit a signal of the form $WB.\mathrm{cosec}\, W.\cos Vt$.

The aperture effects in the two systems may therefore be compared as follows: when the width of the scanning spot approaches the wave-length of the shade-distribution or a multiple of this wave-length, then, in the case of an ordinary television system, the amplitude of the transmitted signal, which is a pure alternating current, increases rapidly, whereas in a velocity-modulation system the signal increases in complexity in the sense that its higher-frequency Fourier components become increasingly important constituents of the signal.

---

* T. M. MacRobert, *Spherical Harmonics*, p. 252 (Methuen 1927).

† *Bell System Technical Journal*, **6**, 602 (1927).

## DISCUSSION

Mr J. H. AWBERY. I find it of distinct interest to notice how frequently methods and results in pure mathematics are applicable in quite unrelated fields of physics. Thus the author's equation (3) is identical, except for the sign of the quantity $C$, with Kepler's equation* in astronomy, which connects the mean anomaly with the excentric anomaly of a planet. In that case, the reversion of the equation is the author's equation (8) with the factors $(-1)^n$ omitted, corresponding to the fact that $J_n(-nC) = (-)^n J_n(nC)$.

The equation is particularly interesting, because it led Bessel to the discovery of his functions†, and also because it was the first transcendental equation satisfactorily solved by Lagrange, who gave the same solution, but with the Bessel functions† written out as series.

---

* R. S. Ball, *Spherical Astronomy*, p. 156 (London: 1908).

† L. Silberstein, *Synopsis of Applicable Mathematics*, p. 96 (London: 1922).

537·322

# THE LONGITUDINAL THERMOELECTRIC EFFECT: (1) COPPER

By P. C. FENG, B.S. (Tsing Hua)
AND
WILLIAM BAND, M.Sc., Professor of Physics,
Yenching University, Peiping, China

*Received November* 16, 1933, *and in revised form March* 28, 1934. *Read in title May* 18, 1934.

*ABSTRACT.* The e.m.f. produced by asymmetrical temperature-distributions in a copper wire is examined, and found to agree well with the formula

$$\mathbf{F} = t\mathbf{G} + a\,(\mathbf{G}.\mathbf{G})^{\frac{1}{2}}\,\mathbf{G} + b\,(\mathbf{G}.\mathbf{G})\,\mathbf{G},$$

where **F** is the potential-gradient at the point where the temperature-gradient is **G**. The two constants $a$ and $b$ are calculated, and their dependence upon tension in the wire investigated. Their values are given by the formulae

$$a.10^{10} = -15{\cdot}5 + 14{\cdot}2\,W + 25{\cdot}0 \exp(-W),$$

$$b.10^{12} = 47{\cdot}8 - 20{\cdot}5\,W - 50{\cdot}0 \exp(-0{\cdot}45\,W),$$

where $W$ is the load in kilograms on the wire of diameter 0·885 mm.

## § 1. INTRODUCTION

According to the statistical theory of thermoelectric phenomena expounded by Sommerfeld and Frank(1), a conductor in which no electric current flows, but in which a temperature-gradient is maintained, will show a potential-gradient which is proportional to the temperature-gradient. This simple linear relationship is usually accepted as accurately describing the experimental facts. It seems always to be assumed that, no matter what shape the temperature curve may take, if the two ends are at equal temperatures the total Thomson e.m.f. integrates to zero.

This assumption was expressed as the "law of Magnus"(2), which stated that in a circuit composed entirely of one homogeneous metal no e.m.f. could exist, whatever temperature-distribution might be introduced. The fact that an e.m.f. was later discovered in a simple iron wire when asymmetrically heated(3) was explained as due to lack of homogeneity of crystalline structure of the iron.

Benedick's results for copper(4) have not been accepted as a contradiction to the law of Magnus. He used the constriction method to maintain the asymmetrical heating: a sharp point on the end of the heated piece of metal was pressed against the large piece of cold metal; at the point of contact is the very high and quite unmeasurable temperature-gradient. Secondary effects could so easily be introduced by the constriction method that little weight has been attached to his results: see the work of Zernicke (5) and Bernade (6).

Experiments on the manifestly homogeneous metal mercury have so far failed to produce consistent results (4), (5); and the work of Seth, Anand and Anand (7), purporting to show the existence of the e.m.f. in homogeneous copper wires without constrictions, is not quantitative enough to settle the question. These workers did not realize the necessity of an analysis of the temperature-distribution, and so missed the significance of their results altogether.

In the present work we have examined a copper wire again, but with a far more precise analysis of the temperatures. It now appears that the e.m.f. can be regarded as of like nature with the Thomson potential as defined by Sommerfeld and Frank, but slightly generalized.

The potential-gradient is shown to be a function of the temperature-gradient alone, but not a simply linear function. Quadratic and cubic terms are needed, which give finite e.m.f. on integration over an asymmetrical temperature-distribution between equal end temperatures. It may be best to retain the term *Thomson potential-gradient* for the linear factor, and so distinguish between it and the factors responsible for the e.m.f. which arises in asymmetrical states.

A quadratic term was at first thought to be inadmissible, since no e.m.f. is observed in perfectly symmetrical states; but finally we found it possible to introduce the term vectorially so that it gave zero e.m.f. for a symmetrical distribution, as required.

## § 2. APPARATUS AND TECHNIQUE

The apparatus for producing controllable asymmetrical temperature states was made to the design shown in figure 1. At one end is a water-cooling jacket, which abuts immediately upon the slate former of the central heating-coil. At the other end of the central heater is a conical brass piece forming the core of another heater, which can be controlled independently of the central heater. To the back of the

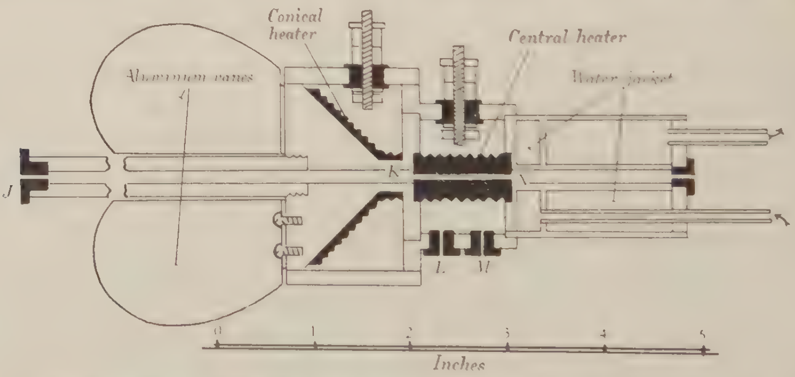


Figure 1. Section through apparatus for producing asymmetrical temperature-distribution. *J*, orifice for temperature probe; *K*, *N*, positions for comparison with figure 2; *L*, *M*, orifices for fixed thermocouple leads.

cone are attached several aluminium radiators and, axially, a long brass rod down the centre of which the specimen wire is carried. Gradual cooling to room-temperature takes place down this narrow tube, and both ends of the specimen are thus maintained at about the same temperature. Since the specimen was a copper wire, copper leads were taken to the potential-measuring instrument.

Two copper-constantan thermocouples are attached permanently to the ends of the central heater to indicate the high temperatures. Another couple is used as a probe from the ends, for exploring the temperature state. A fine copper wire is first wrapped with thin paper, and a fine constantan wire placed in contact with this; the whole is then wrapped again in thin paper; and the ends of each wire project from the sleeve so far that the thermojunction can be made. Provided that temperatures are not taken too high, this kind of probe proves very satisfactory.

To obtain an accurate map of the temperature state, we first get an idea as to where the critical points are situated, the position in the water-cooled section where the temperature suddenly starts to increase, and the point near the maximum where the steepest gradient ends.

The error in temperature caused by the process of probing would be greatest in the steep gradient. For this reason, when the probe thermocouple was inserted through the water-cooled end, we used it only to find the critical point at which the temperature suddenly began to rise. All the other temperatures for which the probe was used were found by inserting it through the other end. Thus the probe did not disturb the steep gradient at all, and the position of the critical point was found to within about 0·2 mm. either way. The steep gradient was found by interpolation between the cold temperature point and the high temperature registered by the first fixed thermocouple in the heater. One source of error lies in the finite size of the fused contact between the couple and the wire, which is nearly 0·5 mm. long. The maximum probable error in our final results due to these various sources is estimated as about 4 per cent.

In finding the small gradient by the probe process we noticed that the thermocouple readings obtained while the probe was gradually inserted through the end differed by no more than 2 or 3 per cent from those obtained as the probe was gradually withdrawn. The enclosure acts as a shield or guard-ring, and the heat-flow is chiefly along the wire. Further, the heat-capacity per cm. of the enclosure is at least 500 times greater than that of either the specimen or the probe. We should thus expect that ideally slow probing would introduce negligible changes in the temperature of the enclosure and of the specimen wire when a steady state is obtained. In practice small changes in the e.m.f. were found to accompany the temperature-variations introduced by the probe, but these errors can be averaged out and are less significant than the errors made in estimating the steep gradient.

For the present work we decided not to attempt the use of a potentiometer; the most sensitive type available here being sufficient only for a minimum reading of 0·5 $\mu$V. Instead we use a high-sensitivity galvanometer* of resistance 15 $\Omega$

---

* No current should be drawn if our results are to be described as e.m.f. measurements, according to Sommerfeld's remarks (1).

manufactured by the Cambridge Instrument Co., Ltd. This meter has a sensitivity of 2 cm. deflection at 1 metre for 1 $\mu$V.

After the e.m.f. for any particular temperature state had reached its steady value, the e.m.f. was recorded between each reading of temperature made during the mapping. This mapping is accomplished by first finding the temperatures of the critical points, and then probing at intervals of 1 or ½ cm. along the rest of the wire.

To eliminate as much as possible of the error thus introduced the temperature map was found by retracing the steps once or twice; the mean curve thus obtained was then used for correlation with the mean value of the e.m.f. found during the whole process.

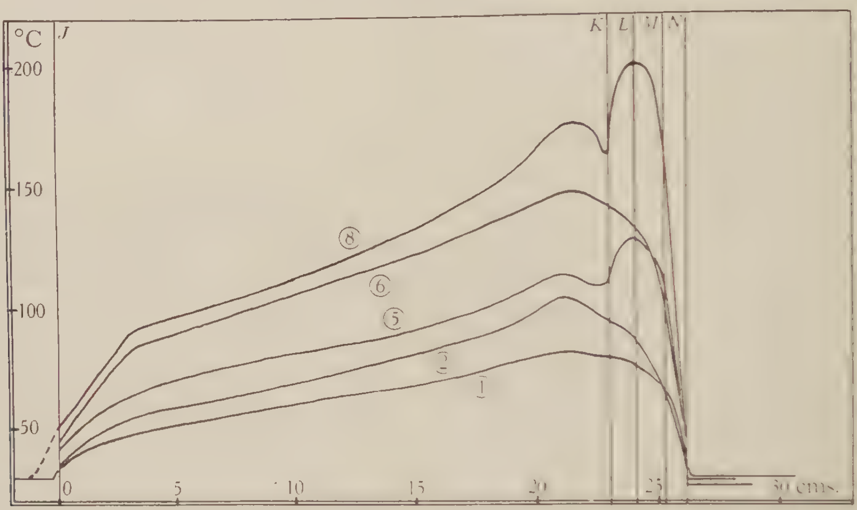


Figure 2. Typical temperature maps. The numbers correspond to those given in § 3. *K*, *L*, *M*, *N*, are positions corresponding to those in figure 1.

A further source of error is that some difficulty was experienced in keeping the heating-currents constant, owing to sudden fluctuations in the mains; but this error must be of less significance than those mentioned already, since continual watch was maintained on the heating-current and its controls.

The only way to increase the precision of our results would be to use smaller temperature-gradients. These could be measured more precisely and directly by means of fixed thermocouples at points near the ends of either gradient. But the smaller e.m.f.'s produced would necessitate a more sensitive galvanometer in order to permit of their estimation with the same degree of precision as the temperature-gradients. With the meter used in the present work higher precision in measuring temperatures would have been useless.

The heating-coils are wound non-inductively, but for the present work on copper this point is probably of no significance; work on iron will however be attempted shortly. Reversal of the heating current had no effect on the e.m.f., which is thus not disturbed by leakage from the heater on to the specimen.

The best temperature curves seem to be those, figure 2, for which the central heater carries the smaller current, for the others show an unfortunate hollow between the two heaters. While the present design gives the possibility of a very wide choice in types of curves, and is thus useful in preliminary work, higher accuracy could be obtained by using only one heater and relying upon shielding-tubes to produce the slow gradient on one side. This we are doing for further work already commenced.

The present apparatus is thus suitable to show that no matter how complicated the temperature state may be we can still correlate the e.m.f. with the temperature-gradient, and thus demonstrate that it is a true Thomson potential in a slightly generalized sense.

No special precautions were taken with regard to the chemical purity of the specimen; the investigation of the effect of impurity and previous mechanical and thermal treatment would make an important field for further study.

## § 3. RESULTS

The results of our measurements are analyzed with the help of the following simple theory. Let the vector electric potential-gradient be **F**, and the vector temperature-gradient be **G**, and assume that $\mathbf{F}$ $\mathbf{G}$

$$\mathbf{F} = t\mathbf{G} + a\,(\mathbf{G}.\mathbf{G})^{\frac{1}{2}}\,\mathbf{G} + b\,(\mathbf{G}.\mathbf{G})\,\mathbf{G} \qquad \text{......(1)},$$

where $t$ is the Thomson potential-gradient of Sommerfeld and Frank, and $a$ and $b$ are new constants to be derived from our results. $t, a, b$

It may be pointed out here that we must use the vector notation in order to include the second term in the expansion (1). A quadratic term in scalar form would give an experimentally incorrect expression, resulting in an e.m.f. under a purely symmetrical temperature-distribution.

To find the total e.m.f. $E$ developed in a wire along which **G** is directed, we have to integrate: $E$

$$\begin{aligned} dE &= \mathbf{F}.d\mathbf{x} \\ &= t\mathbf{G}.d\mathbf{x} + a\,(\mathbf{G}.\mathbf{G})^{\frac{1}{2}}\,\mathbf{G}.d\mathbf{x} + b\,(\mathbf{G}.\mathbf{G})\,\mathbf{G}.d\mathbf{x}, \end{aligned}$$

and since $G = dT/dx$ in magnitude, this may be written

$$E = t\,(T_1 - T_2) + \int a\,(\mathbf{G}.\mathbf{G})^{\frac{1}{2}}\,dT + \int b\,(\mathbf{G}.\mathbf{G})\,dT \qquad \text{......(2)},$$

where $T_1$ and $T_2$ are the end temperatures of the wire.

If the two ends of the wire are at the same temperature, then of course equation (2) shows that the normal Thomson potential-difference vanishes identically. Only if the temperature-distribution is symmetrical, however, will the $a$ and $b$ terms give also a zero result. On division of the curve into suitable sections these two terms can in any particular case be evaluated without difficulty. Thus

$$E = a\Sigma\,(G\,dT) + b\Sigma\,(G^2\,dT) \qquad \text{......(3)},$$

where in the first term the positive magnitude of $G$ must be used without regard to the sign of the gradient.

Altogether eight different temperature states were examined in the first test, of which five are shown in figure 2. The wire was under tension due to application of a weight of 1 kgm. at each end. In each case it was noticed that the positive current flowed from parts with great $G$ to parts with small $G$, in the wire: on open circuit there should therefore be an accumulation of electrons in the specimen at the base of the steeper temperature-gradient.

By grouping the curves into pairs thus: nos. 1 and 5, 2 and 6, 3 and 7, 4 and 8, we calculate the average of the four values thus obtained for the constants $a$ and $b$. The result was:

$$a = \quad 7{\cdot}90\ 10^{-10}\ \text{volt-cm.}/(^\circ\text{C.})^2,$$

$$b = -\ 4{\cdot}65\ 10^{-12}\ \text{volt-cm}^2/(^\circ\text{C.})^3.$$

$I_1$
$I_2$
$E$
$A$
$B$

These values are used to calculate the theoretical e.m.f. in table 1. Here $I_1$ is the current in amperes in the central heater, $I_2$ the current in the conical heater; $E$ is the observed e.m.f. in microvolts averaged over all the values taken during the mapping of the corresponding temperature-distribution; $MD$ is the mean deviation of the individual values of e.m.f. during this process; $A$ is the term $\Sigma\,(G\,dT)$ and $B$ the term $\Sigma\,(G^2dT)$ in equation (3) above. The theoretical value of $E$ is thus found from $E = aA + bB$.

Table 1

| Curve | $I_1$ | $I_2$ | $E$ | $MD$ (per cent) | $A.10^{-2}$ | $B.10^{-4}$ | $aA + bB$ |
|---|---|---|---|---|---|---|---|
| 1 | 2·0 | 4·5 | 0·83 | 4·4 | 13·86 | 5·619 | 0·83 |
| 2 | 1·4 | 6·0 | 1·19 | 3·1 | 19·20 | 7·967 | 1·15 |
| 3 | 2·2 | 5·0 | 1·19 | 2·4 | 20·87 | 8·752 | 1·23 |
| 4 | 2·4 | 5·0 | 1·68 | 1·1 | 28·35 | 13·30 | 1·62 |
| 5 | 3·0 | 6·0 | 2·18 | 0·9 | 42·46 | 25·92 | 2·15 |
| 6 | 2·3 | 7·0 | 2·46 | 2·0 | 53·25 | 39·96 | 2·35 |
| 7 | 1·5 | 7·6 | 2·46 | 1·3 | 55·01 | 39·15 | 2·53 |
| 8 | 3·0 | 8·0 | 2·89 | 1·7 | 173·0 | 240·0 | 2·76 |

The values of $MD$ are unfortunately somewhat large in several cases, and this indicates that the probe method of finding the temperature-distribution had a disturbing effect not altogether negligible.

In order to find, if possible, the dependence of the two constants on the tension, similar work was carried out while the wire was kept under a series of different tensions. The results are given in the graph of figure 3.

In general we can remark that apparently both $a$ and $b$ have non-zero values as the tension approaches zero. But our values for the smallest tension measured are not very reliable, because the weight used ($\frac{1}{2}$ kgm.) was not enough to straighten out the small bends in the wire; and these must introduce local irregularities in the strain of the material, only eliminated when the general tension is large enough to straighten out the wire.

As the tension increases to several thousand kgm. cm$^2$ the curve approaches linearity, both functions increasing with tension much more rapidly than for small tensions.

In figure 3 the full curves are plotted from the formulae

$$a.10^{10} = -15{\cdot}5 + 14{\cdot}2\,W + 25{\cdot}0 \exp(-W),$$
$$b.10^{12} = 47{\cdot}8 - 20{\cdot}5\,W - 50{\cdot}0 \exp(-0{\cdot}45\,W),$$

where $W$ is expressed as kilograms load on the wire. These formulae appear to give about the best fit, but little significance attaches to empirical formulae of the sort unless fortuitously. Thus in this case we have fitted formulae of the forms $a = a_0 - a'(1 - W^2 a''^2)^{\frac{1}{2}}$ and also $a_0 W \exp(a' W^n)$ with nearly equal accuracy, but the latter formula makes the curve pass through the origin and ignores our results for the $\frac{1}{2}$-kgm. load. A simple power series in $W$ might also be found, but we chose the form used above as more directly representing the visible characteristics of the curves.

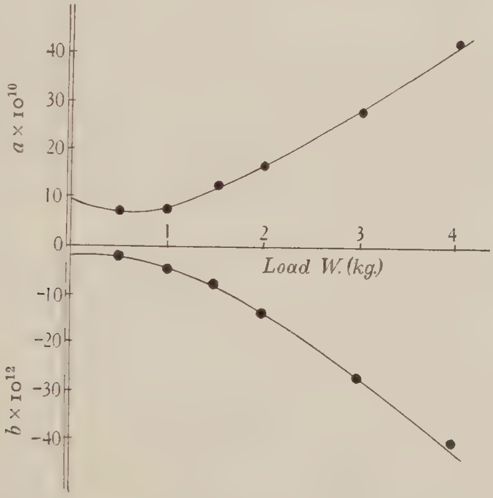


Figure 3. The two new thermoelectric constants as functions of tension. Small circles indicate values calculated directly from observations, and the full curves are plotted from empirical formulae. The values at $W = 1$ are most certain.

We suggest, on the basis of these results, that it would be better to determine more accurately the behaviour of the two constants under small tensions; probably by the use of thin rods suspended vertically, instead of wires. Since, however, the constants must obviously depend greatly upon previous treatment of the substance this was not possible with our specimen, which was a wire of mean diameter 0·885 mm.

## §4. DISCUSSION OF RESULTS

We indicate first how a formula like (1) can be derived theoretically. In Fermi statistics (8), the thermionic current across the face of a crystal is a function of the form

$$i = A.T^2.e^{-E/T},$$

where $A$ and $E$ are constants, $A$ depending on the number of electrons per cm³, and $E$ upon the surface work function. If two crystals are in close contact in our wire specimen, the current flowing across the interface will be the differential

$$di = 2ATS.dT + AT^2.dS,$$

where the value $dS$ signifies the expansion as a series in $dT$ of the differential of the exponential term. If $E/T$ is large the result is found to be

$$di = -ASE.dt + \tfrac{1}{2}ASE^2.dT^2/T^2 - \tfrac{1}{6}ASE^3.dT^3/T^4 \ldots$$

and here $dT$ is the temperature difference between the adjacent faces. Ordinarily we neglect the terms in $dT^2$ and $dT^3$ in taking differential coefficients, but in this case we see that this may introduce a serious error when $dT$ is not really zero and $E/T$ is large. Here $dT$ will be proportional to $G$, the temperature-gradient in the wire, and so we find a form like that of equation (1).

We have pursued this idea in great detail but found finally that quantitatively the results are many orders too small to account for our observations. This probably means that the physical ideas at the basis of the work are erroneous; we include the foregoing remarks merely to point out the necessity for greater care in differentiating functions like the Fermi distribution function. The correction terms in $G^2$ and $G^3$ for the Thomson potential-gradient may be obtainable from a revision of Sommerfeld's calculations from this point of view.

The present work has been reported in some detail to prepare for a more concise account of future results.

## REFERENCES

(1) *Reviews of Modern Physics*, III (January, 1931).
(2) *Pogg. Ann.* (1851).
(3) Trouton. *Proc. R.S. Dublin* (1886).
(4) Benedicks. *Comptes Rendus* **163**, 751 (1916); **165**, 391 (1917); **169**, 578 (1919). *Ann. d. Phys.* **55**, 1 (1918); **62**, 185 (1920).
(5) Zernicke. *Proc. K. Akad. Amst.* **21**, 1262 (1919).
(6) Bernade. *Phys. Rev.* **18**, 199 (1921).
(7) Seth, Anand and Anand. *Indian J. of Phys.* **3**, 431 (1928).
(8) Darrow. *Reviews of Modern Physics*, I (July, 1929).

546.262:541.5

# THE LINKAGE OF $CO_2$ AND CO

By H. LESSHEIM AND R. SAMUEL,
The Muslim University, Aligarh

*Communicated by W. Jevons, March* 2, 1934. *Read May* 18, 1934.

*ABSTRACT.* New calculations of Adel and Dennison make it possible to show that the ground level of $CO_2$ arises adiabatically from C ($p^4\ {}^3P$) + O ($^3P$) + O ($^3P$), or in other words that the linkage of the carbon atom to either of the oxygen atoms is exactly the same and is due to a pronounced $p^2{=}p^4$ bond (O=C=O). Excitation in the ground level of the CO molecule of those two electrons which in the adiabatic dissociation become the $s^2$ group of carbon in such a way that in the dissociation they become two additional $p$ electrons, brings about the carbonyl term of CO. Adiabatic dissociation of $CO_2$ into CO and O leads to the carbonyl term (=C=O) and an unexcited O. The linkage in CO has to be considered as a double link. The numerical agreement is very good.

In a recent paper Adel and Dennison (1) have calculated the constants of the $CO_2$ molecule using the data of the infra-red spectrum. This enables us to determine the energy of adiabatic dissociation of carbon dioxide in a way which was not previously possible. In a recent note (2) on the dissociation of $PbO_2$ the corresponding figures for $CO_2$ were not given, since at the time of writing they could not be compared with other values found by an independent method. It was shown that the combination of a Pb atom with two O atoms arises neither from the ground level of Pb ($s^2p^2\ {}^3P$) nor from any of the levels of the configuration ($sp^3$), but from the anomalous $^3P$ state with four equivalent $p$ electrons, and it was mentioned that the linkage in $CO_2$ is probably of the same type.

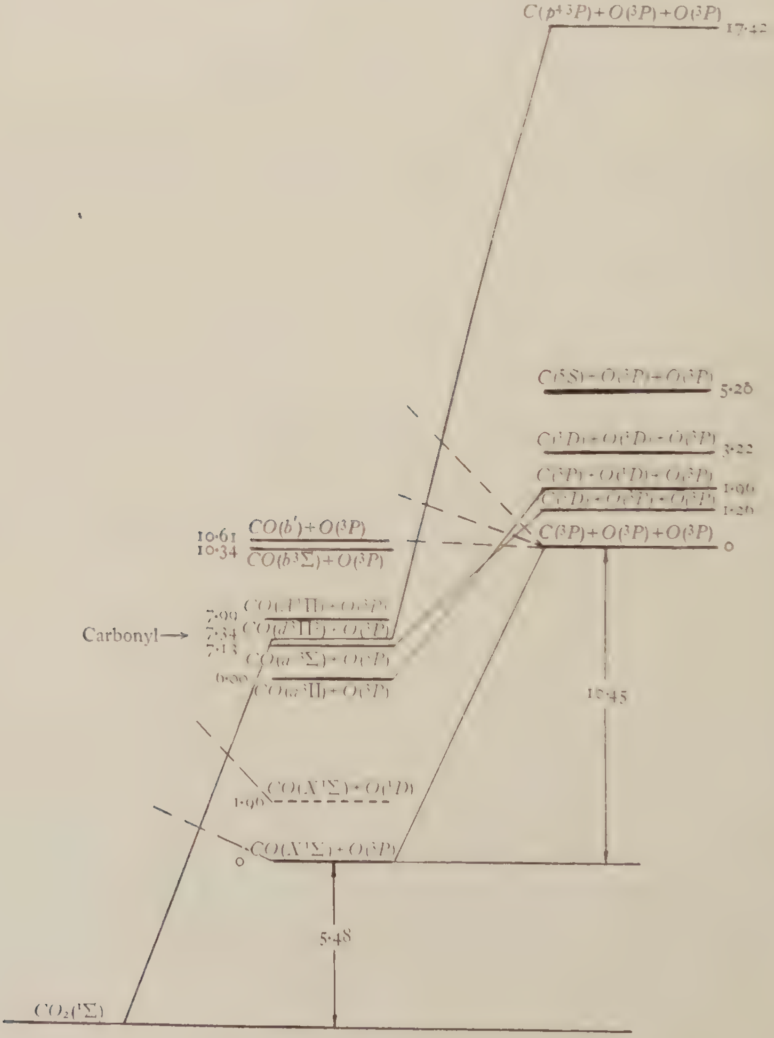
The heat of formation of $CO_2$ from CO and $O_2$ is 67·62 k.-cal./mol. (3), the best value obtainable for the heat of dissociation of $O_2$ is 117·4 k.-cal./mol. Hence

$$\begin{array}{l}
(CO) + \tfrac{1}{2}(O_2) = (CO_2) + \phantom{1}67{\cdot}6 \text{ k.-cal./mol.}\\
(O) \phantom{C}- \tfrac{1}{2}(O_2) = \phantom{(CO_2) + 1}58{\cdot}7 \text{ k.-cal./mol.}\\
\hline
(CO) + \phantom{1}O\phantom{)} = (CO_2) + 126{\cdot}3 \text{ k.-cal./mol.}
\end{array}$$

i.e. the ground level of the $CO_2$ molecule lies 126·3 k.-cal./mol. or 5·48 electron-volts below the ground level of the CO molecule and a separated O atom.

Carbon monoxide in its ground level dissociates into unexcited atoms. The extrapolation of the vibrational levels of the $^1\Sigma$ state gives a value for the dissociation energy which is clearly too high in the light of thermochemical data. Kaplan (4) has suggested the possibility of obtaining the dissociation energy from the levels b $^3\Sigma$ at 10·34 and b′ at 10·61 V. These levels are sometimes considered to be not two distinct electronic levels at all, but consecutive vibrational levels of the same state b $^3\Sigma$. This in our opinion is highly improbable, as b′ does not combine with the ground state whereas b $^3\Sigma$ does, both in emission and in absorption. We can, therefore, say that the second vibrational level of b $^3\Sigma$, which might also lie at about 10·6 V., is already quenched by predissociation. The repulsive state to which the predissociation is due will originate at about 10·45 V., because it will be of a very flat character since it is clearly the lowest of the repulsive states. A still lower repulsive state, if it were capable of existence, would have quenched the first level at 10·34 V.

The value of 10·45 V. obtained in this way is also in good agreement with thermochemical data, if account is taken of the fact that these are subject to a certain error owing to the fact that the heat of sublimation of carbon is not well known for low temperatures. On the other hand our computation allows us to calculate this heat of sublimation spectroscopically. Denoting the heat of sub-
x limation of carbon ($\beta$-graphite) by $x$, we have at 25° C.



$$\begin{array}{lll}
[C]_{\beta\text{-graphite}} + \tfrac{1}{2}(O_2) & = (CO) + & 26{\cdot}65 \text{ k.-cal./mol.} \\
(C) - [C]_{\beta\text{-graphite}} & = & x \\
(O) \qquad - \tfrac{1}{2}(O_2) & = & 58{\cdot}70 \text{ k.-cal./mol.} \\
\hline
(C) \qquad + \quad (O) & = (CO) + & (85{\cdot}35 + x) \text{ k.-cal./mol.}
\end{array}$$

Since

$$(C) \qquad + \quad (O) = (CO) + 241{\cdot}05 \text{ k.-cal./mol. } (= 10{\cdot}45 \text{ V.}),$$

for the heat of sublimation of carbon ($\beta$-graphite) at 25° C. we obtain

$$x = 155{\cdot}7 \text{ k.-cal./mol.}$$

We can now return to the question of the adiabatic dissociation of $CO_2$. The carbon term $p^4\,{}^3P$ is not measured, but can well be computed from the corresponding terms of the ionized atom in the usual manner*. Thus

$$\begin{aligned}
C\,(2s\,2p^3\,{}^5S) - C\,(2s^2\,2p^2\,{}^3P) &= \ \ 5{\cdot}20 \text{ V.},\\
C\,(2p^4\,3P) \quad - C\,(2s\,2p^3\,5S) &= 12{\cdot}22 \text{ V.}
\end{aligned}$$

The total energy of dissociation of $CO_2$ into $C\,(p^4\,{}^3P) + O\,({}^3P) + O\,({}^3P)$ is therefore $5{\cdot}48 + 10{\cdot}45 + 5{\cdot}20 + 12{\cdot}22 = 33{\cdot}35$ V.

Adel and Dennison⁽¹⁾ have obtained theoretical expressions for the energies etc., and calculated the occurring constants in accordance with the infra-red spectrum. Thus for the energy of dissociation of $CO_2$ into CO and O they get the value

$$103750 \text{ cm.}^{-1} \text{ or } 12{\cdot}80 \text{ V.}$$

It can be deduced from the data given in that paper that the energy of dissociation of the CO molecule in that excited state in which it is left when $CO_2$ is stripped of its first O atom or, so to speak, is in the state in which it is able to receive a second O atom, is

$$166350 \text{ cm.}^{-1} \text{ or } 20{\cdot}53 \text{ V.}$$

Hence the total energy of dissociation of $CO_2$ obtained from the infra-red spectrum is

$$270100 \text{ cm.}^{-1} \text{ or } 33{\cdot}33 \text{ V.},$$

and is in surprisingly good agreement with the value of 33·35 V., calculated above. The difference of 7·73 V. between the energies necessary for the dissociation of the first and second O atoms is due to the repulsive force between the O atoms. Now it can definitely be said that tetravalent carbon, in $CO_2$ at least, possesses four $p$ electrons, the two $s$ electrons of divalent carbon having been promoted to the $p$ shell.

Since the ground level of $CO_2$ is a ${}^1\Sigma$ state and the O atoms are linked in their ${}^3P$ states, removal of the first O atom leaves CO in a triplet state; most probably in a ${}^3\Pi$ state. This is certainly not the ${}^3\Pi$ state at 5·98 V., which was taken into consideration by Herzberg(6)[*] and which dissociates into atoms of lower energy of excitation. The triplet state, we expect, must lie 12·80 V. above the $CO_2$ ground state, i.e. 7·32 V. above the ground-most state of CO, having an energy of dissociation of 20·53 V. We suppose that the state termed $d\,^3\Pi$ is the one in question[†].

---

* In current literature the state $sp^3\,{}^5S$ of C is sometimes supposed to be about 1·6 V. above the ground level. This value was assumed by Heitler and Herzberg (1929) in order to explain the intersection of potential curves of CN. The spectrum of $C^+$, however, shows that the proper value is 5·20 V. (See also J. H. Van Vleck, who gives 5 to 8 V⁽¹⁶⁾.) In this particular case owing to the greater number of $p$ electrons the computation from the spectrum of the ion will yield even a better value than in other cases where it has been shown to give fairly exact results⁽⁵⁾.

This triplet term which we have now obtained, whether it is identical with the $d\,^3\Pi$ state or not, is formed by tetravalent carbon with four equivalent *p* electrons and accordingly has two free *p* valencies. It is therefore the state from which compounds of the type $\begin{matrix}R\\R\end{matrix}C{=}O$ arise. CO in this state is the so-called carbonyl (=C=O). Only in the case where R is H is there no difficulty in using a carbonyl arising from C ($sp^3$). To our mind the above calculation definitely establishes that, at least in the state in question, the bond of radicles to the carbonyl is a true *p*—*p* bond and that the possibility of a *q* bond suggested by Hund(7) does not occur here; probably because the levels of CO arising from C ($sp^3$) and C ($p^4$) respectively are not sufficiently close, and do not approach each other either when radicles or a second O atom are brought near.

The fact that the *adiabatic* dissociation of the carbonyl leads us to a carbon atom in the $p^4\,^3P$ state does not of course imply the assumption that the *chemical* dissociation proceeds in this manner. In speaking of an adiabatic dissociation we necessarily mean the continuation of the curve which would be the only one capable of existence if no perturbation occurred. This does not mean that by a non-radiating transition the molecule could not change to an intersecting curve, especially in cases where the multiplicities of the states concerned are equal; this change, therefore, not only is most probable but even might proceed along a slightly varied connexion, thus avoiding the point of intersection. In the example discussed in this paper there are clearly enough intersections of the adiabatic potential curve of the carbonyl state with other potential curves of the same multiplicity, even with such as go to the ground levels of the atoms, and the *chemical* dissociation therefore may easily end at C ($s^2\,p^2\,^3P$) + O ($p^4\,^3P$).

The CO molecule is sometimes said to be triply linked and written as $C\equiv O$. This assumption, as far as band spectroscopy is concerned, is based on a theory developed by Herzberg and Mulliken. The two groups of electrons determining the bond in the ground state of CO are $\pi^1\,(2p)\,\sigma^2\,(2p)$, which doubtless are bonding in all the existing theories. The difference between the theories in this particular case amounts only to a different definition of valency. Herzberg and Mulliken count these six electrons as three pairs and in this way define a trivalency of the CO bond, no matter from which atom the single electrons originate. Hund(7), on the other hand, counts only those pairs in which electrons from different atoms join, and in a recent paper(8) along with considerations about the part played by disturbing electrons the authors have upheld the same point of view. This means that, since two of the six *p* electrons in CO come from C and four from O, we have only two bonding pairs, and accordingly divalency obtains.

---

[*] The states to which Herzberg mainly attributes the $CO_2$ linkage are CO ($^1\Sigma$) and O ($^1D$), to which we shall refer below.

[†] We are indebted to Dr R. K. Asundi for informing us that on the long-wave side of the $d\,^3\Pi$ level now recorded at 7·9 V. he has found several new heads belonging to this level, and that this fact is likely to bring $r_0$ down to our value. This level is obtained under special conditions only in the presence of a rare gas, which may be due to the excitation of two electrons.

The difference between these two views is really only a difference in definition, and so it is impossible to prove by mathematical reasoning that either of them is incorrect. We can only ask which definition agrees better with the usual conception of valency. The definition which is wanted by chemistry is actually a definition of bonds *between the atoms*. The definition given by Mulliken and Herzberg, although it frequently leads to the same result as the other, certainly counts the number of bonds between the six *p* electrons without consideration as to whether these electrons go to the same atom or not in the dissociation. This, however, is not the question with which we are concerned. We wish to obtain information regarding the forces by which the nuclei are kept together. For our purpose the forces between the four *p* electrons of O are immaterial, and when we only consider the interaction between the C atom on one side and the O atom on the other side we come to divalency. According to the definition of Mulliken and Herzberg the same kind of trivalency would always prevail whenever six *p* electrons meet, not only in $N_2$ and CO, but also in BF and even in a BeNe; but it would be a trivalency which keeps only the six electrons of the *p* shell together, whereas the number of bonds between the nuclei of the partners are 3, 2, 1 and 0 respectively. Therefore only that definition which identifies the number of links with the number of pairs of electrons from different partners meets the requirements of chemistry.

It has sometimes been suggested ⁽⁹⁾ that the value 10·45 V. of the dissociation energy of CO points towards a triple linkage on the basis of an average value of approximately 3 V. per link. It is intended to deal with this question at greater length in connexion with general problems of valency; so attention may be confined at this stage to mentioning only that certain regularities can also be produced without Herzberg and Mulliken's method of counting the valencies. In the oxides of the lighter elements 5 V. per link appears to be a truer value than 3 V. if in the case of perturbing electrons we deduct an appropriate amount for perturbation. Thus the dissociation energy is 4·7 V. for OH, 5·75 V. for BeO which is singly linked⁽⁸⁾, and 10·45 V. for CO which is doubly linked. Again, in NO the dissociation energy is 6·7 V. so long as the linkage is perturbed but increases to 10·67 V. as soon as the perturbing electron is excited⁽⁸⁾. We cannot therefore expect $N_2$ to fall into line with the oxides and to show a dissociation energy of 15 V. for the triple bond because of the lower electronic affinity of N. A theory which implies a constant value for the stability of the bond, irrespective of the atoms concerned, cannot be physically sound, since stability must depend on the charges of the nuclei, or (what is the same thing) on ionization potential and electronic affinity. Already rough calculation indicates that we cannot compare a double and a triple link in this manner. If only for the purpose of rough calculation and without any great assumption of accuracy we set the energy of the $\sigma^2$ group equal to half the energy of the $\pi^4$ group, we can say that the splitting of the $\sigma^2$ group and the $\pi^4$ group needs six-ninths or eight-ninths of the total energy according to whether we split the $\pi^4$ group into $1 + 3$ or $2 + 2$ electrons, since in the first case three and in the second case five of the six mutual interactions in $\pi^4$ undergo fission. That means that even when the conditions are not altered by any change in the nuclear charges the transition from a double to a triple link would increase the energy of dissociation not by 50 per cent but only by about 33 per cent.

Herzberg and Mulliken's enumeration would certainly be correct, if the above compounds dissociated adiabatically with always the same distribution of the *p* electrons, i.e. $N_2$ into N – N, CO into the ions $C^- - O^+$, and BF into $B^{--} - F^{++}$; and even a combination of $Be^{---}$ and $Ne^{+++}$ would not only be possible, but also give rise to a triply covalent BeNe. In the case of the CO molecule it is a well-established fact that CO in its ground state dissociates into unexcited atoms, and in the ordinary adiabatic way and without intersection of an ionic curve⁽¹⁰⁾*. Also in chemical literature a triple link is sometimes ascribed to the molecule of carbon monoxide⁽¹¹⁾. Three arguments are given, viz. the proportionately high frequency in the Raman effect, the difference of the dipole moments of carbon monoxide and the carbonyl radicle, and the preservation of the octet rule. The frequency of the Raman line depends on the energy of dissociation and the anharmonic constant of the vibration. If we neglected the possible difference of anharmonicity, a high Raman frequency would denote only a high energy of dissociation, and we have already pointed out above that the dissociation energy of 10·45 V. is by no means an exceptionally high value among those of the oxides of the light elements, and that equal energies of dissociation do not mean an equal number of bonds. Since, however, the Raman frequency is a function also of the anharmonicity, its value gives still less evidence of a triple link. Nor can the dipole moment be taken as a measure of the number of bonds. In view of the electronic configuration of CO in its $^1\Sigma$ state it is only surprising that a dipole moment is present at all in the molecule, indicating a slight polarization of the molecule in spite of its covalent linkage. We have shown above that the carbonyl radicle arises from the carbon monoxide by excitation of two electrons, or perhaps of one in exceptional cases. The symmetrical electronic configuration of carbon monoxide is thus perturbed, and it would be most amazing if a considerable dipole moment did not arise from this perturbation, especially in those cases in which the two former *s* electrons of C then forming an $s^2$ shell of spherical symmetry are now literally bound to the side opposite to the O atom. These two tests failing, there is no physical reason whatever to assume a particular exception for the linkage of the CO molecule for the sole purpose of the preservation of an orthodox octet rule. It may be added that chemical evidence also does not appear to favour the preservation of the octet rule⁽¹²⁾.

Whether the CO linkage is considered to be double or triple, it is clear from this discussion that it is brought about by the two *p* electrons of unexcited carbon. The importance of the conclusions drawn here from Adel and Dennison's calculations lies in the proof that the C atom in $CO_2$ possesses definitely four equivalent *p* electrons. Adiabatically stripping it of an O atom leads us to an excited CO molecule, to be called the carbonyl state of CO, in which C and O are linked by two *p* electrons of carbon whilst two more *p* electrons of the C atom are open to other linkages, which then become true *p*—*p* bonds.

---

* Pauling on his assumption of 3 V. per linkage has attempted to show that CO in its ground state is actually bound as $C^-$ $O^+$. His proof is not conclusive since wave mechanics yields no possibility of exchanging part of the energy curve of one molecular state for that of another. If two potential curves intersect, perturbations occur only at those vibrational levels at which the intersection takes place and in their vicinity, i.e. at those levels for which the energies of the states are exactly or nearly equal.

In former discussions it was sometimes assumed (13) that the CO state responsible for either $CO_2$ or for the carbonyl or for both of them arises from C ($sp^3$), which we see now is only possible in compounds in which an H atom is linked to the C atom. Another assumption is that $CO_2$ arises by combination of CO in its ground state with O in its $^1D$ state, and that the carbonyl radicle is a CO molecule in one of the ordinary excited triplet states of which a $^3\Pi$ is the lowest* (6), (14), (15). This is not only inconsistent with the new results but also open to other objections. CO ($^1\Sigma$) + O ($^1D$) would probably give rise to repulsive states only and would not be compatible with the equivalency of the two O atoms which is established by both optical and chemical evidence. Furthermore, the ordinary excited triplet states of CO, which would according to Mecke be the carbonyl states, arise from the ground state when one of the bonding groups undergoes fission, which in any case means the loss of a link. Even if we call the CO linkage a triple link, the newly arising two valencies would have to be located either on the side of O only or one at O and one at C. If, however, we term the CO linkage a divalency, which we have explained to be more sound, a carbonyl with a single link between carbon and oxygen, and the free valencies again located one at carbon and one at oxygen thus (—C—O—), would remain, and this is not the chemical carbonyl group. It is quite true that the electrons of the same group in the molecule are equivalent and that on account of their exchange a particular electron cannot be ascribed to a particular nucleus; but apart from other reasons already mentioned, the use of structural formulae implies that at any moment in each term a distinct number of electrons belong to a certain atom, especially as far as electrons corresponding to free valencies are concerned.

Mecke has attempted to explain the difference between the thermochemical energies of dissociation of $CO_2$ and CO which seems contradictory to the equivalence of the two oxygen atoms in $CO_2$. On the assumption that the products of dissociation of $CO_2$ into CO and O are normal CO and an excited O atom, he has found a value of about 5 or 15 k.-cal./mol. for the repulsive energy of the O atoms according to the excited term of O he has taken into account. This figure is in complete disagreement with the 7·73 V., equalling about 178·3 k.-cal./mol., which has been calculated by Adel and Dennison (1). The above discussion throws light on the question why the thermochemical energies of dissociation seem to be different. It is only with the help of the value of 7·73 V. for the energy of repulsion, without which the carbonyl state at 7·34 V. would lie exactly half way between CO ($^1\Sigma$) and C ($p^4\ ^3P$) + O ($^3P$) + O ($^3P$), and on the consideration that the actual energies of the links to either O atom are strictly equal, that we arrive at a complete explanation of the facts observed.

---

* This state seems to dissociate into C ($s^2p^2\ ^3P$) and O ($p^4\ ^1D$) since the extrapolated energy of dissociation indicates an atom excited by about 2 V. above the level of 10·45 V.

## REFERENCES

(1) A. Adel and D. M. Dennison. *Phys. Rev.* **44**, 99 (1933).
(2) H. Lessheim and R. Samuel. *Current Science*, **1**, 374 (1933).
(3) Landolt-Börnstein. *Tabellen* (2nd supplement), **2**, 1501.
(4) J. Kaplan. *Phys. Rev.* **37**, 1406 (1931).
(5) W. Grotrian. *Graphische Darstellung der Spektren* (Berlin, 1928).
(6) G. Herzberg. *Z. f. Phys. Chem.* (B), **17**, 68 (1932).
(7) F. Hund. *Z. f. Phys.* **73**, 1 (1931).
(8) H. Lessheim and R. Samuel. *Z. f. Phys.* **84**, 637 (1933).
(9) L. Pauling. *J. Amer. Chem. Soc.* **54**, 3570 (1932).
(10) L. Pauling. *J. Amer. Chem. Soc.* **54**, 988 (1932).
(11) Hammick, New, Sidgwick and Sutton. *J. Chem. Soc.* 1876 (1930).
(12) R. F. Hunter and R. Samuel. *J. Chem. Soc.* (in the press).
(13) R. Mecke. *Z. f. Phys. Chem.* (B), **7**, 108 (1930).
(14) R. Mecke. *Z. f. Phys. Chem.* (B), **18**, 53 (1932).
(15) R. Mecke. *Trans. Far. Soc.* **30**, 200 (1934).
(16) J. H. Van Vleck, *J. Chem. Phys.* **2**, 20 (1934).

535.338.1:546.14.02

# THE SPECTRUM OF TREBLY-IONIZED BROMINE, Br IV

By A. S. RAO, M.A., D.Sc.,

AND

S. GOPALA KRISHNAMURTY, M.A., Department of Physics,
Andhra University, Waltair, India

*Communicated by K. R. Rao, D.Sc., January* 10, 1934.
*Read April* 20, 1934.

*ABSTRACT.* Vacuum spark and discharge-tube spectra of bromine have been investigated over a wide range. Many of the triplets and singlets involving the terms due to $4p$, $5s$, $4d$, $sp^3$ and $5p$ configurations have been identified. From the present work the classifications made by S. C. Deb appear to be incorrect. The deepest term $4p\,^3P_0$, which equals 404892 cm.$^{-1}$, yields an ionization potential of about 50 volts for the trebly-ionized atom of bromine.

## § 1. INTRODUCTION

Our knowledge of the spectrum of Br IV consists chiefly of the list of wave-lengths of lines attributed to Br IV by L. and E. Bloch(1) from experiments on the electrodeless discharge through bromine vapour, in the comparatively short region $\lambda 3261$ to $\lambda 2302$. The only attempt at the analysis of this spectrum has been that of S. C. Deb(2) who arranged some of the lines into multiplets arising from the transition $5p \rightarrow 5s$ of the series electron. From the present investigation, however, the classification suggested by Deb appears to be erroneous.

In continuation of the work on the spectra of Br V, VI and VII(3), a detailed investigation of the spectrum of Br IV has been undertaken. It is the purpose of the present paper to describe the results of the analysis of the lines due to the trebly-ionized atom of bromine. The experimental work which forms the basis of the present paper has been fully described in the previous paper on Br V, VI and VII. It may, however, be stated that in the quartz region many new lines have been recorded in addition to the list published by Bloch. By a careful study of the plates in the extreme ultraviolet region, taken under different degrees of excitation and covering the range $\lambda 430$ to $\lambda 1500$, the lines of Br IV have been selected.

## § 2. STRUCTURE OF THE SPECTRUM

The spectrum of Br IV is analogous to the iso-electronic spectra of Ge I, As II and Se III, the essential features of which are almost completely known(4). This has considerably facilitated the analysis of the spectrum of Br IV. The theoretical term scheme of Br IV is given in table 1.

Table 1. Predicted terms of Br IV

| $3_1$ | $3_2$ | $3_3$ | $4_1$ | $4_2$ | $4_3$ | $4_4$ | $5_1$ | $5_2$ | $5_3$ | Term prefix | Terms |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 6 | 10 | 2 | 2 | . | . | . | . | . | $4p$ | $^3P\ ^1D\ ^1S$ |
| 2 | 6 | 10 | 2 | 1 | . | . | 1 | . | . | $5s$ | $^3P\ ^1P$ |
| 2 | 6 | 10 | 2 | 1 | 1 | . | . | . | . | $4d$ | $^3F\ ^3D\ ^3P\ ^1F\ ^1D\ ^1P$ |
| 2 | 6 | 10 | 2 | 1 | . | 1 | . | . | . | $4f$ | $^3G\ ^3F\ ^3D\ ^1G\ ^1F\ ^1D$ |
| 2 | 6 | 10 | 2 | 1 | . | . | . | 1 | . | $5p$ | $^3D\ ^3P\ ^3S\ ^1D\ ^1P\ ^1S$ |
| 2 | 6 | 10 | 1 | 3 | . | . | . | . | . | $sp^3$ | $^5S\ ^3D\ ^3P\ ^3S\ ^1D\ ^1P$ |

The first evidence of the regularities in the spectrum of Br IV has been the discovery of the recurrence of the characteristic interval 2990 cm.$^{-1}$ among several lines in the region about λ570, which have been definitely ascribed to the trebly-ionized atom after a careful scrutiny of the plates. Some of the strongest lines in the spectrum are to be sought in the transition $5s-4p$ of the series electron. By the application of the irregular doublet law to the lines due to the combination between the deep terms, it is expected that the chief triplet group $4p\,^3P-5s\,^3P$ lies in the region λ550. In table 2 the position of the line $4p\,^3P_2-5s\,^3P_2$ in the analogous spectra of Ge I, As II, Se III is shown, and the corresponding line identified in Br IV is included also.

Table 2.

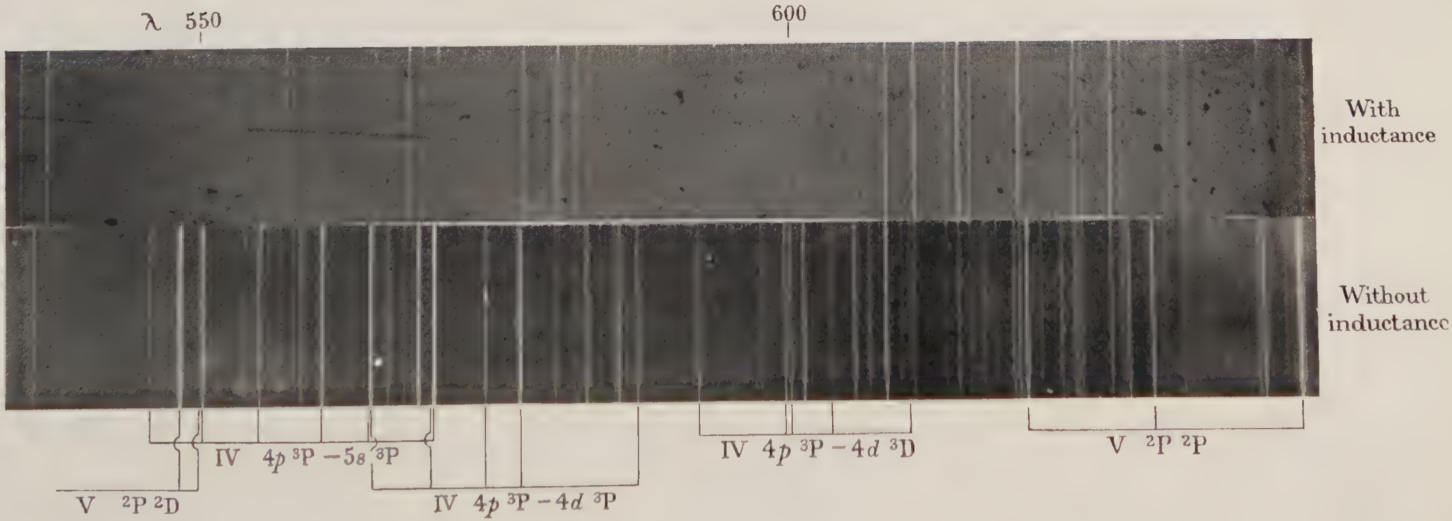
| Spectrum | $4p\,^3P_2-5s\,^3P_2$ | $\nu'=\nu-2468(Z-A)^2$ | $\delta\nu'$ |
|---|---|---|---|
| Ge I | 37708 | 35240 | |
| As II | 78967 | 69095 | 33855 |
| Se III | 126454 | 104242 | 35147 |
| Br IV | 180350 | 140862 | 36620 |

It is found that some of the above-mentioned pairs occur just in this region. A further comparison of these spectra made by the application of the regular doublet law to the intervals of the $4p\,^3P_1-4p\,^3P_2$ in the sequence Ge I to Br IV has clearly indicated that the difference 2990 cm.$^{-1}$ must be identified as the interval $4p\,^3P_1-4p\,^3P_2$. The screening-constant for Br IV on this assumption is shown in table 3.

Table 3. Regular doublet law

| Element | $4p\,^3P_1-4p\,^3P_2$ | $\sqrt[4]{(\delta\nu/0\cdot0455)}$ | $s$ | $\delta s$ |
|---|---|---|---|---|
| Ge I | 853 | 11·69 | 20·31 | |
| As II | 1477 | 13·42 | 19·58 | 0·73 |
| Se III | 2196 | 14·82 | 19·18 | 0·40 |
| Br IV | 2990 | 16·01 | 18·99 | 0·19 |

Spark spectrum of rubidium bromide.



*To face page* 533.

Table 4 gives the combinations identified in the present work. The starting-point for the classification is the identification of the ground multiplet $4p\,^3P-5s\,^3P$. The irregular doublet sequence, the progression of the screening-constant, and the uniform behaviour of the lines forming the multiplet (shown in plate 1) constitute the chief evidence for the correctness of the identification. The identification of this ground multiplet receives additional support by the location, in the quartz region, of the super-multiplet due to the transition $5p \rightarrow 5s$ of the running electron.

Table 4. Combinations in Br IV

| | | $4p\,^3P_0$<br>404890 | $4p\,^3P_1$<br>401643 | $4p\,^3P_2$<br>398653 | $4p\,^1D_2$<br>386775 |
|---|---|---|---|---|---|
| | | | *3247* | *2990* | |
| $5s\,^3P_0$ | 224271 | | 177372 (9) | | |
| | *1281* | | | | |
| $^3P_1$ | 222990 | 181900 (9) | 178653 (9) | 175663 (6) | — |
| | *4688* | | | | |
| $^3P_2$ | 218302 | | 183341 (5) | 180350 (8) | — |
| $5s\,^1P_1$ | 215856 | — | 185787 (2) | — | 170918 (9) |
| $4d\,^3D_1$ | 235980 | 168910 (9) | 165659 (5) | — | 150787 (2) |
| | *1115* | | | | |
| $^3D_2$ | 234865 | | 166778 (8) | 163786 (2) | 151912 (1) |
| | *2856* | | | | |
| $^3D_3$ | 232009 | | | 166644 (8) | 154762 (1) |
| $4d\,^3P_0$ | 225937 | | 175706 (10) | | |
| | *1625* | | | | |
| $^3P_1$ | 224312 | (180578) | 177331 (8) | 174340 (8) | — |
| | *− 3892* | | | | |
| $^3P_2$ | 228204 | | 173440 (10) | 170447 (10) | — |
| $sp^3\,^3D_1$ | 250993 | 153898 (6) | 150648 (10) | 147658 (5) | — |
| | *8053* | | | | |
| $^3D_2$ | 242940 | | 158701 (10) | 155710 (10) | — |
| | *10608* | | | | |
| $^3D_3$ | 232332 | | | 166318 (10) | — |
| $sp^3\,^3P_0$ | 236555 | | 165088 (5) | | |
| | *− 348* | | | | |
| $^3P_1$ | 236903 | (167987) | 164740 (6) | 161748 (5) | — |
| | *− 419* | | | | |
| $^3P_2$ | 237322 | | 164322 (8) | 161328 (7) | — |
| A | 265711 | 139179 (10) | 135933 (9) | — | — |
| B | 226897 | — | 174746 (3) | 171757 (6) | 159872 (4) |
| C | 222411 | — | 179230 (6) | 176244 (5) | — |

Supermultiplet due to the transition $5p \rightarrow 5s$

| | | $5p\,^3P_0$<br>188103·6 | $5p\,^3P_1$<br>185427·0 | $5p\,^3P_2$<br>182859·8 | $5p\,^1D_2$<br>177127·0 |
|---|---|---|---|---|---|
| | | | *2676·6* | *2567·2* | |
| $5s\,^3P_0$ | 224271 | | 38843·2 (3) | | |
| $^3P_1$ | 222990 | 34886·4 (2) | 37563·0 (4) | 40130·1 (5) | 45862·9 (3) |
| $^3P_2$ | 218302 | | 32872·4 (3) | 35439·6 (5) | — |
| $5s\,^1P_1$ | 215856 | — | — | — | 38730·2 (7) |

According to the scheme proposed by S. C. Deb (2) the intervals of the $5s\,^3P$ term are $5s\,^3P_0 - 5s\,^3P_1 = 506$; and $5s\,^3P_1 - 5s\,^3P_2 = 1093$ so that $5s\,^3P_0 - 5s\,^3P_2 = 1599$ cm.$^{-1}$, a value less even than that of Ge I. Such an abnormally low value appears to be improbable from the following table in which the intervals $5s\,^3P_0 - 5s\,^3P_2$ of Ge I-like spectra are given together with the screening constants.

Table 5

| Element | $5s\,(^3P_0 - {}^3P_2) = \delta\nu$ | $Z - s$ | $s$ |
|---|---|---|---|
| Ge I | 1669 | 16·36 | 15·64 |
| As II | 2777 | 18·58 | 14·42 |
| Se III | 4114 | 20·49 | 13·51 |
| Br IV (writers) | 5970 | 22·49 | 12·51 |
| (Deb) | 1599 | 16·18 | 18·82 |

Perhaps it might be suggested that the regular doublet law is often found to be inapplicable to the intervals of terms other than those of the deepest ($4p\,^3P$ in this case). But it would be difficult to assume that such abnormality occurs suddenly at Br IV. There is yet another interesting sequence which confirms the value of 5970 cm.$^{-1}$ for the interval $5s\,^3P_0 - 5s\,^3P_2$ of Br IV. According to Goudsmit's theory of multiplet separations this interval should be slightly less than the limit $mp\,^2P_{\frac{1}{2}} - mp\,^2P_{\frac{3}{2}}$ of the next higher ion, since the levels $5s\,^3P_{0,1,2}$ and $5s\,^1P_1$ are derived by the addition of an $s$ electron to the ion in the $^2P$ state. This rule is found to hold good remarkably well in almost all the spectra(5) of the type under consideration, as is shown in table 6.

Table 6

| | C I | N II | O III | F IV | Si I | P II | S III | Cl IV |
|---|---|---|---|---|---|---|---|---|
| $ms\,(^3P_0 - {}^3P_2)$ | 60·1 | 167·96 | 375·3 | 618 | 271·93 | 528·1 | 449·7 | 1446 |
| $(m+1)\,s\,(^3P_0 - {}^3P_2)$ | — | 170·70 | 373 | — | 282·9 | 546·8 | 924·2 | — |
| $(m+2)\,s\,(^3P_0 - {}^3P_2)$ | — | 173·4 | — | — | 284·4 | — | — | — |
| Lt. $mp\,(^2P_{\frac{1}{2}} - {}^2P_{\frac{3}{2}})$ | C II | N III | O IV | F V | Si II | P III | S IV | Cl V |
| | 64·0 | 174·4 | 387 | — | 287 | 566 | 950 | 1492 |

| | Ge I | As II | Se III | Br IV |
|---|---|---|---|---|
| $ms\,(^3P_0 - {}^3P_2)$ | 1666·1 | 2778 | 4114·1 | 5970 |
| $(m+1)\,s\,(^3P_0 - {}^3P_2)$ | 1740·2 | — | 4354·3 | — |
| $(m+2)\,s\,(^3P_0 - {}^3P_2)$ | 1764·9 | — | — | — |
| Lt. $mp\,(^2P_{\frac{1}{2}} - {}^2P_{\frac{3}{2}})$ | Ge II | As III | Se IV | Br V |
| | 1768 | 2940 | 4376 | 6085 |

The analysis of Br IV suggested by Deb, which is essentially based on the identification of the interval 1599 cm.$^{-1}$ for $5s\,^3P_0 - 5s\,^3P_2$, appears therefore to be erroneous from the above considerations. A further test of the correctness of the identification will be afforded by the location of the remaining levels of the $5p$ configuration, namely $^3D$, $^3S$, etc.

With a knowledge of the characteristic intervals of the $4p\,^3P$ term, obtained by the writers, it has been possible to locate, though only partially, the terms due to the $4d$ and the $sp^3$ configurations of the atom. All the combinations so far obtained are shown in table 4. There can be no doubt as to the reality of these levels that have been discovered for they have been based on very careful measurement of the lines, which could be confidently assigned to Br IV. Owing to the abnormally large values of the separations involved, however, the allocation of the levels to the various terms of the respective states of the atom cannot be considered final.

It may be pointed out that the $4d\,^3D$ term is regular in the spectra of As II and Br IV while in Ge I and Se III it shows partial inversion. The absence of the combination $4p\,^3P_0-4d\,^3P_1$ may throw some doubt on the identification of the $4d\,^3P_1$ level; but it is significant that this combination is missing in Se III also. The lines $\nu\nu$ 175706, 166318 and 165088 did not form pairs giving any characteristic separations, and hence the assignment of these lines to the combinations suggested in the above scheme is to be regarded as tentative.

## § 3. TERM VALUES

Term values of Br IV are calculated from the two members of the $mp\,^3P$ series that have been identified in the present work thus:

$$5s\,^3P_2-4p\,^3P_2=-\ 180350,$$

$$5s\,^3P_2-5p\,^3P_2=35439{\cdot}6;\ \text{from which}\ 5s\,^3P_2=218302,$$

and the deepest term $4p\,^3P_0 = 404890\ \mathrm{cm}^{-1}$ yielding a value of about 50·0 volts for the ionization potential of the trebly-ionized bromine atom. The terms and their values are listed in table 7.

Table 7. Term values of Br IV

| Term | Term value | Term | Term value |
|---|---|---|---|
| $4p\,^3P_0$ | 404890 | $4d\,^3P_2$ | 228204 |
| $4p\,^3P_1$ | 401643 | $b$ | 226897 |
| $4p\,^3P_2$ | 398653 | $4d\,^3P_0$ | 225937 |
| $4p\,^1D_2$ | 386775 | $4d\,^3P_1$ | 224312 |
| $a$ | 265711 | $5s\,^3P_0$ | 224271 |
| $sp^3\,^3D_1$ | 250993 | $5s\,^3P_1$ | 222990 |
| $sp^3\,^3D_2$ | 242940 | $c$ | 222411 |
| $sp^3\,^3P_2$ | 237322 | $5s\,^3P_2$ | 218302 |
| $sp^3\,^3P_1$ | 236903 | $5s\,^1P_1$ | 215856 |
| $sp^3\,^3P_0$ | 236555 | $5p\,^3P_0$ | 188103·6 |
| $4d\,^3D_1$ | 235980 | $5p\,^3P_1$ | 185427·0 |
| $4d\,^3D_2$ | 234865 | $5p\,^3P_2$ | 182859·8 |
| $sp^3\,^3D_3$ | 232332 | $5p\,^1D_2$ | 177127·0 |
| $4d\,^3D_3$ | 232009 | | |

In table 8 a list of all the classified lines, together with the classifications suggested in the present work, is given.

Table 8. Classified lines of Br IV

| $\lambda$ | $\nu_{vac.}$ | Classification | $\lambda_{vac.}$ | $\nu_{vac.}$ | Classification |
|---|---|---|---|---|---|
| 3041·18 (3) | 32872·4 | $5s\,{}^3P_2-5p\,{}^3P_1$ | 607·02 (6) | 164740 | $4p\,{}^3P_1-sp^3\,{}^3P_1$ |
| 2865·61 (2) | 34886·4 | $5s\,{}^3P_1-5p\,{}^3P_0$ | 605·74 (5) | 165088 | $4p\,{}^3P_1-sp^3\,{}^3P_0$ |
| 2820·87 (5) | 35439·6 | $5s\,{}^3P_2-5p\,{}^3P_2$ | 603·65 (5) | 165659 | $4p\,{}^3P_1-4d\,{}^3D_1$ |
| 2661·40 (4) | 37563·0 | $5s\,{}^3P_1-5p\,{}^3P_1$ | 601·26 (10) | 166318 | $4p\,{}^3P_2-sp^3\,{}^3D_3$ |
| 2581·19 (7) | 38730·2 | $5s\,{}^1P_1-5p\,{}^1D_2$ | 600·08 (8) | 166644 | $4p\,{}^3P_2-4d\,{}^3D_3$ |
| 2573·68 (3) | 38843·2 | $5s\,{}^3P_0-5p\,{}^3P_1$ | 599·60 (8) | 166778 | $4p\,{}^3P_1-4d\,{}^3D_2$ |
| 2491·14 (5) | 40130·1 | $5s\,{}^3P_1-5p\,{}^3P_2$ | 592·03 (9) | 168910 | $4p\,{}^3P_0-4d\,{}^3D_1$ |
| 2179·77 (3) | 45861·9 | $5s\,{}^3P_1-5p\,{}^1D_2$* | 586·69 (10) | 170447 | $4p\,{}^3P_2-4d\,{}^3P_2$ |
| $\lambda_{vac.}$ | $\nu_{vac.}$ | | 585·08 (9) | 170918 | $4p\,{}^1D_2-5s\,{}^1P_1$ |
| 735·66 (10) | 135933 | $4p\,{}^3P_1-a$ | 582·22 (6) | 171757 | $4p\,{}^3P_2-b$ |
| 718·50 (10) | 139179 | $4p\,{}^3P_0-a$ | 576·57 (10) | 173440 | $4p\,{}^3P_2-4d\,{}^3P_1$ |
| 677·24 (5) | 147658 | $4p\,{}^3P_2-sp^3\,{}^3D_1$† | 573·59 (8) | 174340 | $4p\,{}^3P_2-4d\,{}^3P_1$ |
| 663·80 (10) | 150648 | $4p\,{}^3P_1-sp^3\,{}^3D_1$ | 572·26 (3) | 174746 | $4p\,{}^3P_1-b$ |
| 663·19 (2) | 150787 | $4p\,{}^1D_2-4d\,{}^3D_1$ | 569·27 (6) | 175663 | $4p\,{}^3P_2-5s\,{}^3P_1$ |
| 658·28 (1) | 151912 | $4p\,{}^1D_2-4d\,{}^3D_2$ | 569·13 (10) | 175706 | $4p\,{}^3P_1-4d\,{}^3P_0$ |
| 649·78 (6) | 153898 | $4p\,{}^3P_0-sp^3\,{}^3D_1$ | 567·40 (5) | 176244 | $4p\,{}^3P_2-c$ |
| 646·15 (1) | 154762 | $4p\,{}^1D_2-4d\,{}^3D_3$ | 563·92 (8) | 177331 | $4p\,{}^3P_1-4d\,{}^3P_1$ |
| 642·22 (10) | 155710 | $4p\,{}^3P_2-sp^3\,{}^3D_2$ | 563·79 (9) | 177372 | $4p\,{}^3P_1-5s\,{}^3P_0$ |
| 630·12 (10) | 158701 | $4p\,{}^3P_1-sp^3\,{}^3D_2$ | 559·74 (9) | 178653 | $4p\,{}^3P_1-5s\,{}^3P_1$ |
| 625·50 (4) | 159872 | $4p\,{}^1D_2-b$ | 557·94 (6) | 179230 | $4p\,{}^3P_1-c$ |
| 619·86 (7) | 161328 | $4p\,{}^3P_2-sp^3\,{}^3P_2$ | 554·48 (8) | 180350 | $4p\,{}^3P_2-5s\,{}^3P_2$ |
| 618·25 (5) | 161748 | $4p\,{}^3P_2-sp^3\,{}^3P_1$ | 549·75 (9) | 181900 | $4p\,{}^3P_0-5s\,{}^3P_1$ |
| 610·55 (2) | 163786 | $4p\,{}^3P_2-4d\,{}^3D_2$ | 545·43 (5) | 183341 | $4p\,{}^3P_1-5s\,{}^3P_2$ |
| 608·56 (8) | 164322 | $4p\,{}^3P_1-sp^3\,{}^3P_2$ | 538·25 (2) | 185787 | $4p\,{}^3P_1-5s\,{}^1P_1$ |

* From Vaudet's list(6).

† Coincidence with low stage line.

## § 4. ACKNOWLEDGMENT

It is a pleasure to thank Dr K. R. Rao for kindly placing his plates in the Schumann region taken with the Seigbahn vacuum spectrograph at our disposal and for many helpful suggestions during the course of the work.

## REFERENCES

(1) L. and E. Bloch. *Ann. d. Phys.* **7**, 205 (1927).
(2) S. C. Deb, *Proc. R. S.* A. **127**, 208 (1930).
(3) A. S. Rao and K. R. Rao, *Nature*, **131**, 170 (1933); A. S. Rao and K. R. Rao, *Proc. Phys. Soc.* **46**, 163 (1934).
(4) K. R. Rao, *Proc. R. S.* **124**, 472 (1929); A. S. Rao, *Proc. Phys. Soc.* **44**, 343 (1932) and *Indian J. Phys.* 7, 561 (1933); J. S. Badami and K. R. Rao, *Proc. R. S.* **140**, 387 (1933).
(5) Cf. K. R. Rao, *Proc. R. S.* **124**, 472 (1929); and Ingram, *Phys. Rev.* **33**, 907 (1929).
(6) Vaudet, *Comptes Rend.* **185**, 1272 (1927).

536.21

# A METHOD FOR THE MEASUREMENT OF THE THERMAL DIFFUSIVITY OF POOR CONDUCTORS

By E. D. van REST, B.A., B.Sc.,
Forest Products Research Laboratory

*Received October* 20, 1933, *and in revised form January* 22, 1934. *Read May* 18, 1934.

*ABSTRACT.* A new method of determining the thermal diffusivity of poor conductors is described. An irregularly shaped block may be used provided that it has a fairly large plane surface. Temperature-observations are made at two points within the block either by embedding thermocouples or by inserting them between cut surfaces. Small temperature-intervals (less than 10° C.) are used and considerable liberty of experimental conditions is afforded. Only a distance and quantities proportional to the temperature-changes at the two points need be measured. The method has been developed for use with wood, but results are given for a block of cast sulphur and several other common materials.

## § 1. INTRODUCTION

A METHOD was required for determining for wood the thermal diffusivity or the constant $h^2$ in the diffusion equation

$$\delta\theta/\delta t = h^2 . \delta^2\theta/\delta x^2.$$

$\theta, t, h, x$

Thermal diffusivity is an important property of wood, more important perhaps than its conductivity, since wood is often used to insulate against temperature-changes rather than against a steady flow of heat. Moreover, the coefficient of diffusivity is used in calculations involved in such processes as seasoning, sterilization against beetle attack, impregnation and the glueing of veneers and plywood.

The method had to satisfy certain requirements introduced by the nature of the material ultimately to be investigated. The chief of these were (1) that the duration of the experiment should be short in order that movement of moisture, if any occurred, should not be appreciable; (2) that moderately large test pieces should be employed in order that the data obtained might be useful in practice, for determinations of the thermal conductivity of poor conductors such as wood have often involved thicknesses small enough to be comparable with the coarse structure of the specimens; and (3) that only small intervals of temperature should be used, since a fairly large variation of the coefficient with temperature was expected.

The Fourier diffusion equation has been applied to many particular cases, most of which have, at some time or other, formed the bases of experimental methods for determining either conductivity or diffusivity. The case of the semi-infinite solid with one plane boundary seemed the most easily realized and was therefore adopted.

Reference to literature shows that this case has been used by Kirchhoff and Hansemann*, Schulze† and Gruneisen‡, all of whom worked on metals.

A great disadvantage of the method used by Kirchhoff and Hansemann and by Schulze was the assumption that the plane boundary was instantaneously brought to the temperature of a stream of water passing over it. Their measurements also depended on an exact knowledge of the time at which this occurred. Gruneisen evaded these assumptions by fitting his readings to theoretical ones by means of a previously calculated table. Although Gruneisen claims that his method does not require a knowledge of the temperature of the boundary, this table is calculated by means of an arbitrary constant which is in fact proportional to the temperature of the boundary. His paper does not indicate how this constant is chosen.

The present method differs from Gruneisen's in providing a graphical means of rapidly carrying out the fitting process, including a unique choice of the figure corresponding to Gruneisen's arbitrary constant. In addition Gruneisen's presentation of the theory is simplified. The method possesses, in common with Gruneisen's, the following advantages: (1) It does not necessitate a knowledge of the temperature of the heated face; (2) it does not necessitate exact registration of the zero of time; (3) only the distances apart of the temperature-measuring instruments are required, not their distances from the plane boundary; (4) quantities proportional to temperature-changes, not temperatures themselves, are measured; (5) the temperature-interval used is small (less than 10° C.).

Because of the inhomogeneity of wood and because precise measurements of its diffusivity by other methods are not available for comparison, the method has been tested initially on sulphur in the rhombic form. Sulphur has the advantages of being non-hygroscopic, and pure and consequently reproducible; it has, moreover, a diffusivity comparable with that of many woods. Further, values of the diffusivity can be calculated from available data on the conductivity, density and specific heat.

## § 2. THEORY OF THE METHOD

$x$ The block extends to infinity in all directions except the negative $x$ direction. The plane $x = 0$ is the heated face. The uniform initial temperature of the block is taken as the zero of measurement.

The solution of the Fourier diffusion equation

$$\delta\theta/\delta t = h^2 \delta^2\theta/\delta x^2 \qquad \text{......(1),}$$

with the conditions

$$\theta = 0, \text{ when } t = 0,$$

$\theta_0$

$$\theta = \theta_0, \text{ when } x = 0,$$

is

$$\theta = \frac{2\theta_0}{\sqrt{\pi}} \int_q^\infty e^{-\beta^2} d\beta \qquad \text{......(2)§,}$$

$q$ where $q = x/2h\sqrt{t}$; $t$ is time and $h^2$ the diffusivity.

---

[*] *Wied. Ann.* **9**, 1 (1880).
[†] *Ibid.* **66**, 207 (1898).
[‡] *Ann. d. Phys.* **3**, 43 (1900).
[§] See, for instance, Ingersoll and Zobel, *Mathematical Theory of Heat Conduction*, p. 78.

Now suppose that a temperature-rise is observed at $x_1$ which can be represented by

$$\theta = \frac{2\theta_0}{\sqrt{\pi}} \int_{q_1}^{\infty} e^{-\beta^2} d\beta \qquad \text{......(3)},$$

where $q_1 = x_1 + S/2h\sqrt{(t+T)}$.

$q_1$, $x_1$, $S$

Then to get an expression for the temperature rise at $x_2$ (where $x_2 > x_1$) we require to solve equation (1) with the conditions

$$\left.\begin{aligned} &\theta = 0 \text{ when } t = 0 \\ &\theta = \frac{2\theta_0}{\sqrt{\pi}} \int_{q_1}^{\infty} e^{-\beta^2} d\beta \text{ at } x = x_1 \end{aligned}\right\}.$$

Now $$\theta = \frac{2\theta_0}{\sqrt{\pi}} \int_{q}^{\infty} e^{-\beta^2} d\beta,$$

where $q = (x+S)/2h\sqrt{(t+T)}$ is a solution of equation (1) satisfying these conditions so that it is the solution required for all points such that $x > x_1$. Therefore when $x = x_2$

$x_2$

$$\theta = \frac{2\theta_0}{\sqrt{\pi}} \int_{q_2}^{\infty} e^{-\beta^2} d\beta \qquad \text{......(4)},$$

where $q_2 = (x_2 + S)/2h\sqrt{(t+T)}$.

$q_2$

The three quantities $\theta_0$, $S$, and $T$ allow the fitting of the actual temperature-rise at $x_1$ to a theoretical one of the form required. In effect, they state that temperature-rises at $x_1$ and $x_2$ such as those observed would have been caused by the instantaneous establishment from a time $-T$ of a constant temperature $\theta_0$ at a plane distant $-S$ from the plane $x = 0$.

$T$

Since $$1/q^2 \text{ or } 4h^2(t+T)/(x+S)^2 \qquad \text{......(5)}$$

is a linear function of $t$ for a given value of $x$, straight lines should result if $1/q^2$ be plotted against $t$. The slopes of the lines for $x_1$ and $x_2$ are respectively

$$m_1 = 4h^2/(x_1+S)^2 \text{ and } m_2 = 4h^2/(x_2+S)^2$$

$m_1$, $m_2$

so that $$h^2 = (x_2 - x_1)^2 . (m_1 m_2)/4\,(m_1^{\frac{1}{2}} - m_2^{\frac{1}{2}})^2.$$

For any one experiment only one value of $\theta_0$ will give straight lines for these plots, so that it can be obtained by trial. $S$ and $T$ do not have to be chosen since they are involved in the lines so chosen, $S$ in the slope and $T$ in the intercept on the time axis. Thus the temperature conditions at the plane $x = 0$ are not brought into the calculation at all.

The one essential condition of the experiment is that the temperature-rise at $x_1$ should be represented by a curve of the family of equation (1), a condition which can be closely complied with experimentally in many different ways—as many ways as there are pairs of values of $S$ and $T$. Actually, in the work to be described, it was found sufficient to allow the plane $x = 0$ to heat up under hand control, to a predetermined temperature, the process taking from one-half to two minutes, and then to come under automatic control and retain constant temperature while temperature-readings within the block were taken. A few trials soon indicated a schedule of heating which would satisfy the conditions of the theory.

## § 3. EXPERIMENTAL ARRANGEMENT

In casting the block of sulphur special precautions were necessary in order to avoid the consequences of the shrinkage which takes place on solidification. Electric heating elements on the metal lining to the mould enabled the cooling of the casting to be controlled. The block used measured 5 in. × 5 in. × 5 in. and had four thermocouples cast in position on the axis. The cross-section was so chosen in relation to the depth of the lowest thermocouple that sideways conduction was negligible according to a theoretical investigation by Peirce and Wilson*.

The specific gravity of the casting, as determined from about two cubic inches cut near the thermocouples, was 1·93. The specific gravity of rhombic sulphur is 2·07, but this cannot be attained in a casting since molten sulphur sets first in the monoclinic form (the density of which is 1·96) and afterwards changes slowly into the rhombic form, leaving a slightly porous casting.

Blocks of the other materials used were made with the same dimensions. In order to avoid difficulties of drying, the wooden test pieces were only 1 in. square by 6 in. long but were surrounded by guard rings of the same material. The thermocouples (made of foil and covering the whole cross-section) were inserted by slicing the inner test piece. The block and its heating plate, together with asbestos boards at top and bottom, were then placed in a clamp holding all together with a light pressure. A liberal covering of cotton wool round the sides served to minimize any temperature-fluctuations which might be introduced from the sides.

The leads from the thermocouples were taken to a mercury-cup switch-box and thence to a Tinsley vernier potentiometer. Since, in the use of the potentiometer, transits of the galvanometer-spot across the zero were timed for certain settings of the dials, it was an advantage to keep the resistances of the thermocouple circuits as low as possible. The movement of the spot was then more rapid and the timing of the transits more definite. Over the small range used in any one experiment (less than 10° C.) the variation of thermal e.m.f. with temperature could be taken as linear. However, in order that the initial temperatures might be known, the couples were calibrated at the National Physical Laboratory.

The heating plate consisted of a 400-ohm 220-volt asbestos-woven resistance mat clamped between two 16-s.w.g. sheets of copper. Five copper-eureka thermocouples in series were also placed between the plates, but insulated from the mat, to actuate the control.

Two resistances in series, figure 1, enabled the heating-current to be adjusted, first for rapid initial heating at various oven temperatures, and afterwards for slow heating with about 10 watts while the control was in operation. Experience indicated good settings for each of these conditions.

The control is represented diagrammatically in figure 1. The whole of the controlling relay was actuated from d.-c. 220-V. mains.

---

* *Proc. Am. Acad.* **34**, 1 (1899).

## §4. THE METHOD

The block under test was placed in the oven and raised to the temperature at which its diffusivity was to be measured. The control for the plate was set about 10° C. higher. When the block had attained a uniform temperature, as indicated by the thermocouples, the heating-plate and a stop-watch were switched on. As the plate neared its controlled temperature, indicated by the mirror galvanometer, the external series resistance was smoothly altered until, at the desired controlled temperature, slow heating or cooling took place according to the position of the relay. After an interval of 5 min. from the commencement of heating, temperature-readings were begun. Each couple was connected to the potentiometer in turn and readings were taken at intervals over a period of 20 min. Preliminary work had shown that no appreciable rise of temperature took place at the far end of the block during the total time of the experiment, so that for the purposes of this experiment the block could be considered semi-infinite. As a check on the readings a smooth curve was fitted to the microvolts plotted against time for each couple. These curves also assist in choosing the value of $\theta_0$ which enters into equation (3).

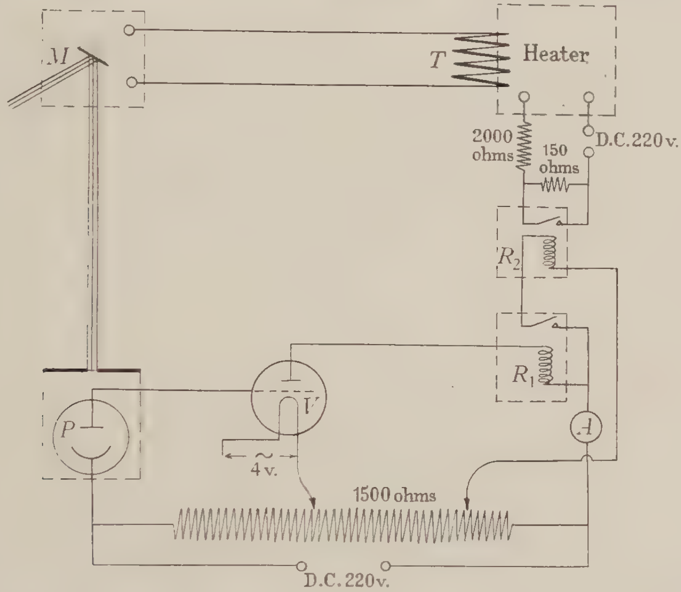


Figure 1. Apparatus for temperature-control of heating-plate. *P*, photocell; *V*, valve (Marconi PX4); *A*, milliammeter; $R_1$, relay (5 mA. Post Office); $R_2$, relay (30 mA.); *M*, mirror-galvanometer (Tinsley).

With a value of $\theta_0$ chosen in this way, $q^{-2}$ was plotted against $t$. The readings of the uppermost couple were plotted on this special scale for several values of $\theta_0$ until the value which gave a straight line for the plot was found. This is not difficult, for, starting with too large a value of $\theta_0$, we obtain a line concave to the time axis; if the chosen value of $\theta_0$ is decreased, the line becomes less concave, then straight, and afterwards more and more convex to the time axis. The choice of values of

*S* and *T* is automatically included, for only one value of $\theta_0$ will give a straight line, and *S* and *T* determine respectively the slope and intercept on the time axis of that line. In practice a very few trials gave the desired value of $\theta_0$. This value was then used to plot the straight lines for the other, lower couples. Any departure of these from straightness indicates that the conditions of the theory have not been exactly fulfilled. All the plotting was done on a drawing-board measuring 30 in. × 24 in. A steel rule divided into tenths of an inch was fastened to one edge and served as the time scale, and a special *T* square gave the scale of $q^{-2}$, equation (5). The divisions on this square are marked with values of $\theta/\theta_0$ from 0 to 0·65, but the
$\Delta$ distance of any mark $\Delta$ from the zero is proportional to $q^{-2}$, where $q$ and $\infty$ are those limits of the definite integral of equation (2) for which it has the value $\Delta$. A table of values of $q$ and $\Delta$ appears in the *Mathematical Theory of Heat Conduction* by Ingersoll and Zobel and, more conveniently with equal intervals of $\Delta$, in *Tables for Statisticians and Biometricians* by Pearson (part II, table II). In the last-named publication, values of the argument require to be multiplied by 2 for the purpose of this paper.

Figure 2 is an example of a set of these straight lines from an actual experiment. The temperature-rise in microvolts for couple 1 is shown in the same diagram.

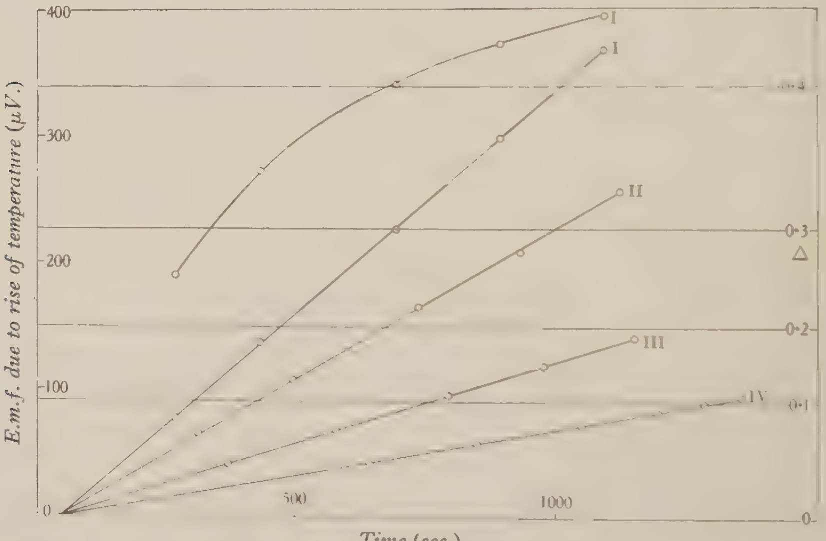


Figure 2. The upper curve shows the potential difference (μV.) corresponding to the actual temperature-rise at couple 1. Straight lines I to IV are for couples 1 to 4 on a special scale.

## § 5. RESULTS

The results of individual experiments at various initial temperatures are set out in table 1. As is usual in determinations of diffusivity, the measurement of distance contributes the greatest risk of error in the result. For this reason the results calculated from couples 1 and 4, involving the largest distance, are included in the table. Only one pair of couples is actually necessary; the others were provided as a check.

Table 1. Thermal diffusivity of rhombic sulphur

(*a*) Calculated from couples 1 and 2. (*b*) From couples 1 and 4.

| Temperature (°C.) | | 29·8 | 27·8 | 26·8 | 28·3 | 27·8 | 28·0 | 75·0 | 75·0 | 75·1 | 74·6 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $h^2 \times 10^4$ (c.g.s. units) | (*a*) | 16·8 | 17·0 | 16·8 | 17·0 | 16·7 | 17·3 | 15·3 | 14·7 | 15·1 | 15·1 |
| | (*b*) | 16·4 | 17·3 | 17·3 | 16·6 | 16·5 | 17·2 | 14·8 | 13·7 | 14·4 | 14·9 |

Inspection of the graphs showed that their results were not so reliable as those from the uppermost pair of couples. In the case of the sulphur block, the measurement of distance was made by means of a travelling microscope. The distance measured was that between the uppermost surfaces of the beads of the couples. There are no determinations of thermal diffusivity of sulphur for comparison, but by using separate values of specific heat and conductivity given by different observers, it is possible to calculate a figure; though, of course, it has the drawback that the constituent determinations were not made on the same material.

In table 2 the observed values of diffusivity are compared with values calculated from conductivities obtained by Kaye and Higgins* and by Green†. The calculation uses specific heats obtained by interpolation from Wigand's‡ results. The density was assumed to be 1·96 in the case of the first-named results; for the others it was known to be 1·94.

Table 2. Comparison with calculated values

(*a*) Kaye and Higgins. (*b*) Green.

| Temperature (° C.) | $k \times 10^3$ (c.g.s. units) | | Specific heat* | $h^2 \times 10^4$ (c.g.s. units) | | |
|---|---|---|---|---|---|---|
| | (*a*) | (*b*) | | Calculated | | Observed† |
| | | | | (*a*) | (*b*) | |
| 20 | 0·650 | 0·615 | 0·1716 | 19·3 | 18·5 | 17·2 |
| 40 | 0·613 | 0·602 | 0·1722 | 18·2 | 18·0 | 16·4 |
| 60 | 0·581 | 0·588 | 0·1731 | 17·1 | 17·5 | 15·6 |
| 80 | 0·558 | 0·570 | 0·1742 | 16·3 | 16·8 | 14·8 |

* Interpolated from Wigand's values.
† Interpolated from mean values of table 1, column (*a*).

The lowness of the observed as compared with the calculated values cannot be attributed to any difference of density, for Green's results show that the density and conductivity of sulphur change approximately in the same proportion, for a small change at least, so that the diffusivity, involving as it does the quotient of conductivity and density, is practically unaltered by a change of density.

The diffusivities of various other substances have been measured by this method but, their constitution being indefinite, comparison with values calculated from published values of density, conductivity and specific heat would not be expected to be very close. The comparison is made in table 3.

---

* *Proc. R.S.* A, **122**, 633 (1929).
† *Proc. Phys. Soc.* **44**, 295 (1932).
‡ *Ann. d. Phys.* **22**, 64 (1907).

Table 3. Thermal diffusivities

| Substance | $k \times 10^3$ (c.g.s. units) | Specific heat | Density | $h^2 \times 10^4$ Calculated | $h^2 \times 10^4$ Observed | Observed density |
|---|---|---|---|---|---|---|
| Oak* | 0·28† | 0·33‡ | 0·60† | **14·1** | **18** | 0·64 |
| Oak | 0·47§ | 0·33 | 0·82§ | **17·4** | | |
| Balsa ¶ | 0·11† | 0·33 | 0·10† | **33·4** | **20** | 0·16 |
| Balsa | 0·13§ | 0·33 | 0·14§ | **28·1** | | |
| Sand, dry | 0·78§ | 0·19** | 1·52§ | **27·0** | **24** | 2·2 |
| Glass, plate | 1·78§ | 0·16** | 2·49§ | **44·7** | **40** | 2·5 |
| Plaster of Paris (cast) | 0·7†† | 0·26‡‡ | 1·2‡‡ | **22·4** | **18** | 1·1 |
| Paraffin wax | 0·60§§ | 0·69** | 0·9** | **9·7** | **6·5** | — |

* Tangential flow, 14 per cent moisture (based on dry weight).
† Griffiths and Kaye. *Proc. R.S.* A, **104**, 73 (1923).
‡ Dunlap. *U.S. Dept. of Agric.*, Bull. No. 110 (1912).
§ *Am. Soc. Ref. Eng. Circ.* No. 1 (1924).
¶ Tangential flow, 13 per cent moisture (based on dry weight).
** Kaye and Laby. *Physical and Chemical Constants.*
†† Lees and Chorlton. *Phil. Mag.* **41**, 495 (1896).
‡‡ *International Critical Tables*, V, 99 and II, 123.
§§ Lees. *Phil. Trans.* **183**, 481 (1892).

## §6. ACKNOWLEDGMENTS

Thanks are due to Sir Ralph Pearson, C.I.E., LL.D., F.L.S., Director of the Forest Products Research Laboratory, for permission to publish this paper, and to Mr W. Barkas, M.Sc., in charge of the Section of Physics, for much helpful advice and criticism.

## DISCUSSION

Dr Ezer Griffiths. I am much interested in the paper, as I have been concerned with the measurement of the diffusivity of the ground. I used an electrically heated sphere buried in the ground with thermocouples located at definite distances from the centre. The heat-supply was so arranged that the temperature of the surface of the sphere followed a sinusoidal curve; this simplified the mathematics required in calculating the diffusivity. The value was checked by independent determinations of the conductivity, the specific heat, and the density, from which the diffusivity was calculated.

I do not think sulphur is a suitable material for the author to use in checking the method as a cast block would have different properties in various directions. Possibly a substance like pitch would be better.

674.03:581.81

# MEASUREMENT OF THE CELL-SPACE RATIO IN WOOD BY A PHOTOELECTRIC METHOD

By WILFRED W. BARKAS, M.Sc., Forest Products Research Laboratory, Princes Risborough

*Received January* 26, 1934. *Read May* 18, 1934.

*ABSTRACT*. A method for measuring the cell-space ratio, or proportion of void space, in wood depends on measuring, by photoelectric means, the amount of light passing through a stained microsection of the wood. A correction for the amount of light penetrating the stained cell-wall is necessary. Comparisons are made between the results obtained by this method and those obtained by the use of a Shand's micrometer, and the conclusion is drawn that the photoelectric method is the more accurate and very much the more rapid. A description of the photoelectric amplifier is given.

## § 1. INTRODUCTION

A KNOWLEDGE of the percentage of void space, or cell-space ratio, of a wood is essential to the solution of certain problems concerning its structural and mechanical properties, and investigations at present proceeding at the Forest Products Research Laboratory have shown the desirability of a simpler and quicker method than the one in general use.

The common method of measuring the cell-space ratio is to traverse the cross-section of the wood under a microscope with a Shand's micrometer, measuring the total distances across the walls and voids separately. Several traverses are made along lines truly representative of the differences in cell-growth, but the figures obtained for the cell-space ratio are only reproducible to within about ± 5 per cent. Moreover, since the diameters of the cells vary from about 0·3 mm. for large vessels down to 0·01 mm. or less for small fibres, the operation is always tedious, especially where large areas are to be measured or where the cells are very small (as for instance in the fibres of ash), and in the latter instance the accuracy must suffer since the width of the image of the micrometer pointer may be comparable with that of the diameters of the smallest fibres.

No comparison seems hitherto to have been made between the results obtained with the Shand's micrometer and those obtained by any independent method, but a paper entitled "Note on the combined use of the photoelectric cell and the projection microscope*," by Savage and Jamieson, gave an indication of the possibilities of using a light-sensitive cell as an alternative method. The authors of that paper measured the average size of certain spores, stained red and mounted in jelly, by projecting a magnified image of a known number of the spores on to a photoelectric cell and comparing the deflection of a sensitive galvanometer with that given by a clear slide. The chief sources of error here lie in assuming that the photocell is insensitive to the light transmitted by the stained bodies, and in the unreliability of measurements made by galvanometer deflections.

---

* *Canadian J. of Res.* 3, 322–6 (1930).

This method was tried at the Forest Products Research Laboratory for the measurement of the cell-space ratio, but it soon became evident that the inherent errors were too large to allow of accurate results and that a correction for the amount of light transmitted by the stain would need to be made.

## § 2. THEORY

For the measurement of cell-space ratio of woods, since the ratio of void area to total area in a transverse microsection is equal to the ratio of void volume to total volume in the material in bulk, a principle similar to that used by Savage and Jamieson in the case of the stained spores may be followed. If a section could be stained absolutely black then the cell-space ratio would be simply the ratio of the light transmitted by the section to the total light transmitted by an equal clear area. In practice a black stain is not obtainable, but if the transmission of the stained wall is known a simple correction may be applied.

$Q_1$, $Q_2$ Let $Q_0$, $Q_1$ and $Q_2$ denote the quantities of light transmitted respectively by equal areas of the clear portions of the slide, of the stained cell walls, and of the
$A$ section as a whole. Then, if we denote the cell-space ratio by $A$, we have

$$Q_2 = AQ_0 + (1 - A) Q_1.$$

Since for a vacuum photocell working at saturation potential the current is proportional to the total quantity of light entering the cell, we may write

$$A = \frac{Q_2 - Q_1}{Q_0 - Q_1} = \frac{C_2 - C_1}{C_0 - C_1},$$

$C_1$, $C_2$ where $C_0$, $C_1$ and $C_2$ are the currents corresponding to the quantities $Q_0$, $Q_1$ and $Q_2$ respectively.

For convenience the values of $Q_1$ and $Q_2$ are expressed throughout the paper as percentages of $Q_0$, and $A$ as a percentage of the area of the section under consideration.

This formula requires a knowledge of the transmission ratio, $Q_1/Q_0$, of the stain on the cell wall. It is not feasible to magnify the image sufficiently for a portion of the cell wall to cover the aperture of the photocell and thus to measure $Q_1/Q_0$ directly, because the light-intensity would be too low. The depth of stain on the cell wall must therefore be matched visually with one of a graded series of filters of the same stain by means of comparison microscopes at equal magnifications, each being illuminated at equal intensity by the light used in the photometer. The transmission $Q_1/Q_0$ of the filter chosen may then be measured on the photometer under the same conditions as is the transmission of the microsections.

## § 3. APPARATUS

(i) *The wood sections* used were about 1 cm. square and cut to a thickness of 15 $\mu$. The use of a thicker section was considered undesirable, because any inaccuracy in cutting perpendicular to the grain meant that light would pass obliquely through the cell spaces so that a portion of the void area would be eclipsed, and the thicker the section the greater would be the amount of light eclipsed. In most woods, too, the cells are not all strictly parallel; for instance, the fibres are slightly inclined in passing round a medullary ray. The sections were stained in a concentrated dye, washed in water, air-dried without heat and mounted in Canada balsam under a cover slip. Since the slide, mounting-medium, and cover slip all absorb or reflect an appreciable amount of light (about 10 to 15 per cent by measurement) the value of $Q_0$ must be measured through these. For this reason a long cover slip was used so as to leave, alongside the section, an equal clear area. The annual rings of the section were placed parallel to the edge of the slide, so that the area included in the rectangular aperture of the photocell would be representative of a definite growth-period in the wood-section.

(ii) *The matching-filters* were made by mounting, on a glass slide, a glass ring of internal diameter 2 cm. and thickness 1 mm. This was filled with an aqueous solution of the dye used as stain, and was closed with a 2·4-cm. cover slip. The whole was sealed with Canada balsam and dried slowly without heat. By varying the concentration of the dye, filters of any desired depth could be made and a graded series of about ten filters, transmitting from about 1 to 15 per cent of light, was prepared.

(iii) *Optical arrangement*, figure 1. The slide $S$ carrying the microsection was mounted on the rising and racking front of a camera body from which the lens had been removed*. An image of $S$ was focused by an f/2·4 camera lens $L_3$ on to a rectangular aperture $A$ placed immediately in front of the photocell $P$. The shape of the aperture could be varied to suit the area under measurement. To enable the focus to be judged and the area for examination to be chosen, a mirror $M$ was so fixed as to rotate about $C$ into the light-path and thus turn the image on to the ground-glass screen $G$ on which the area of $A$ was marked out. In this way any portion of the section might be chosen. When the mirror was out of action it served to close the back of the screen $G$. The two portions between the section $S$ and the mirror $M$ were enclosed in bellows. By suitably spacing $L_3$ and $S$, any magnification between 1·4 and 8 could be obtained.

(iv) *Comparison microscopes.* For matching the depth of stain of the section with one of a graded series of filters, the comparison microscopes ($M_1$ and $M_2$, figure 1) were set up and illuminated from $B$ by the prisms $P_1$ and $P_2$ and a suitable system of lenses. The substage condensers of the comparison microscopes were removed and the light was concentrated in two nearly parallel beams by a system of lenses. These lenses were adjusted until the intensities of the two fields were equal when viewed, at the same magnification, with a clear portion of a slide on one microscope stage and a clear filter on the other. If the condensers are not removed, matching becomes impossible because the light passes through the slide in a flat cone and suffers a certain degree of concentration by refraction. Under a magnification of about 500 the colour of the cell walls was clearly seen and the matching filter could be chosen with considerable accuracy provided the staining was dark. If the stain is too light, different portions of the cell wall stain to different depths, and, further, for light stains the eye is much less sensitive than for deep stains to small differences in transmission between one filter and another. It has been found that the cell wall of most woods may be stained with methylene blue to pass only about 5 per cent of the sodium light. Above about 12 per cent the matching becomes unreliable for the reasons mentioned above.

---

* A rotating mount for the slide, such that the growth rings in a crooked section could be aligned with the edge of the aperture $A$, would be an improvement.

## §4. DEVELOPMENT OF METHOD

To test for the best illuminating colour and the best stain, the optical system in figure 1 was modified in such a way as to enable any narrow band of wave-lengths from the continuous spectrum of a 12-V. line-filament car lamp to be passed through a fine slit and used for illuminating the slide at *S*, the colour band being selected by rotating the prism through which the light passed.

For the purpose in view it was more important that the colour effects should be tested under the working conditions of figure 1 than that the actual values of the wave-lengths employed or their real intensities should be known. For this reason the spectrum was plotted out in terms of the rotation of the prism, while the response of the photocell to any colour was taken as a percentage of the maximum response recorded.

The response of the photocell to all parts of the continuous spectrum showed that this maximum response lay just in the infra-red and fell off rapidly towards the violet; the extreme infra-red was absorbed by the glass. In order that the photocell should not respond to invisible rays, which the eye could not match in the filters, a 2-inch glass cooling-tank was introduced and curves were taken for varying concentrations of $CuSO_4$ solution as percentages of the clear deflections. From this a $\frac{1}{2}$ to 1-per-cent concentration was chosen as the best for absorbing the infra-red without unduly reducing the intensity in the visible region.

The response of the cell to the continuous spectrum was seen at once to be very different from that of the eye, and in the yellow-to-red region the sensitivity of the cell increases, while that of the eye decreases, with increase in wave-length. Thus it was suspected that the visual matching of filters to the stained sections would be unreliable if even a slight change in the colour of the dye were to occur on staining the section, since this change might be in regions of the spectrum which, though affecting the photocell, would not be noticed by the eye. Trials showed this to be the case. Even when a very narrow band of colours in the orange was used both for matching and for measurement of transmission, the range of colour was sufficient to introduce these errors.

From this it was concluded that only a monochromatic source could, with safety, be employed in this method. Use was therefore made of the new Zeiss sodium-vapour lamp, while a methylene blue stain was chosen for the sections, since measurement showed that it possessed a strong absorption band in the region of the *D* lines. Since the intensity of other lines (e.g. of helium) in the spectrum of the lamp is very low, the prism and slit arrangement could be dispensed with, the final apparatus then taking the form shown in figure 1. The 1 per cent $CuSO_4$ cooling cell may be retained as a precaution, though the heat radiation from this lamp is much weaker than that from a filament bulb.

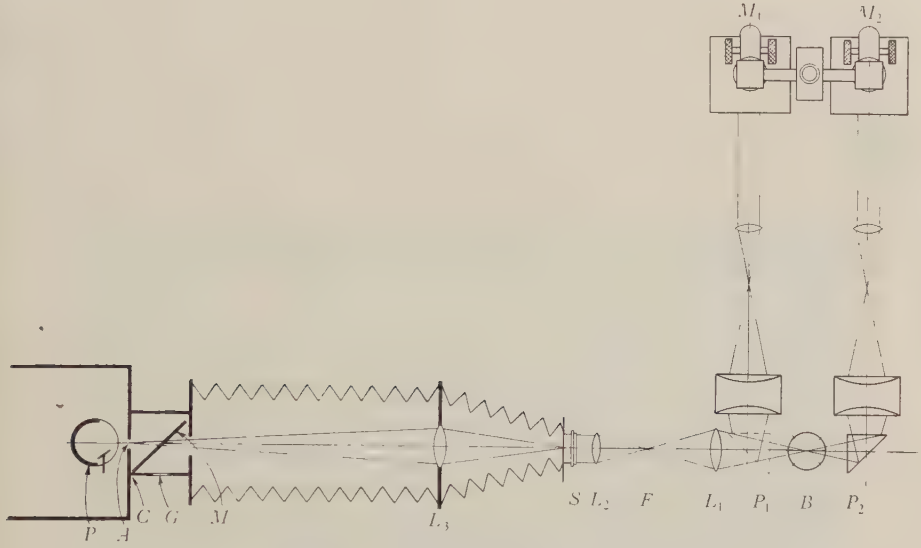


Figure 1. Final lay-out of apparatus.

## § 5. TEST OF METHOD

Since other uses to which the amplifier had been put* had shown that this part of the apparatus is reliable, provided that a constant source of illumination is used, no special tests were made upon it. Attention was therefore confined to tests of the method (i) for accuracy where personal judgment is not necessary in the choice of matching filter; (ii) for reproducibility of measurements on several microsections of identical cell-space ratio but different depths of stain; (iii) for agreement in the choice of matching filter made by several observers on the same microsections; (iv) for the significance of the differences obtained on microsections of different cell-space ratio. These tests are detailed below:

(i) First, to test the method in a case where the continuous filter is a perfect match, grids were cut from translucent papers such as are used for dark-room lanterns. The bars and clear spaces were each as nearly as possible 1 mm. wide.

---

* W. W. Barkas, "Notes on an optical method of measuring 'finish' on a planed wood surface," unpublished report, Forest Products Research Laboratory (1933).

The matching filter was, in this case, a continuous portion of the same material alongside the grid. The grids, after being mounted on a glass slide, were measured under a travelling microscope to determine the percentage of void space. Table 1 shows the results obtained with the photometer, at a low magnification, when an equal number of bars and spaces were included in the field of view:

Table 1. Percentage of void area of paper grid

| Colour of grid | Light passed by | | Void area, as measured by | |
|---|---|---|---|---|
| | paper | grid | photo-meter | direct measurement |
| Green | 1·0 | 49 | 48·5 | 49·6 |
| Red | 7·0 | 55 | 50·7 | 50·9 |

This showed that the method is sufficiently accurate in the case of an ideally matched filter.

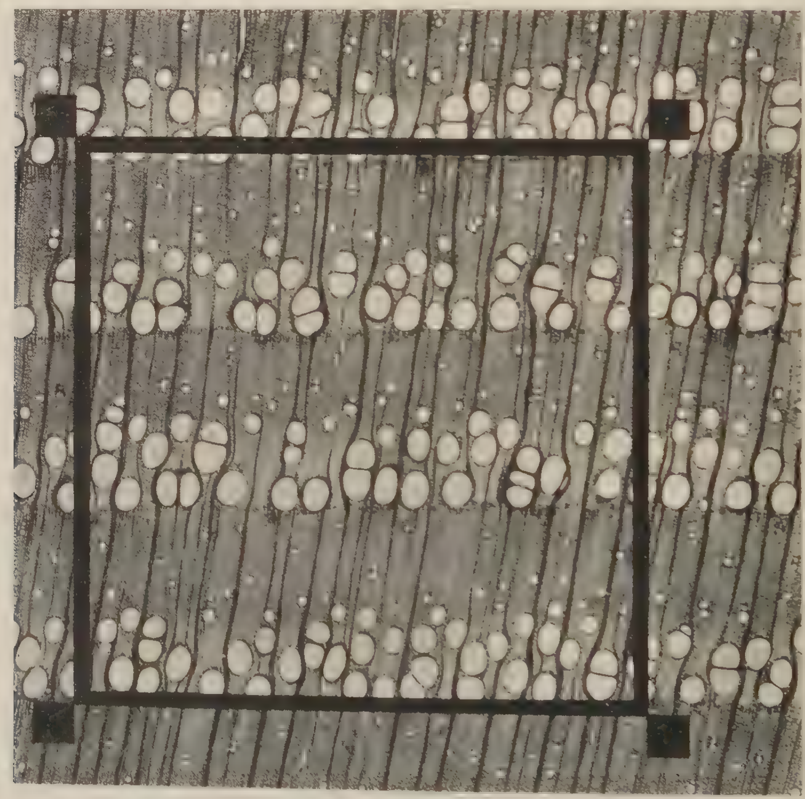


Figure 2. Test section of ash showing area measured (magnification 17.6).

The possibility that the light in passing through the fine holes in the section may be scattered outside the aperture of the photocell is not tested by this method.

Since no method suggested itself whereby a structure of *known* void area and translucence, having voids comparable in size with those of a wood section, might be constructed, the only check on the existence of such an error appeared to be the comparison of the value of the cell-space ratio obtained by means of the photometer with that given by some other independent method. This is dealt with in § 6.

(ii) Next, to test the accuracy of results obtained by matching, a set of 10 sections, 15 $\mu$. thick, were cut successively from a sample of ash chosen for straight grain and even annual rings, figure 2. The total thickness of wood removed from the block was thus only 0·015 cm., so that no appreciable differences were to be expected between the average cell-space ratios of the individual sections. These sections were given identical treatment in mounting except for the depth of stain. The same square area, the side of which was exactly equal to the width of 3 annual rings, was chosen on each section, figure 2, and the magnification was set so that exactly this area and no more was included in the field of view. The photometer measurements were made immediately after matching to avoid possible errors due to fading.

Table 2. Comparison of cell-space ratios obtained in the case of identical sections having different depths of stain

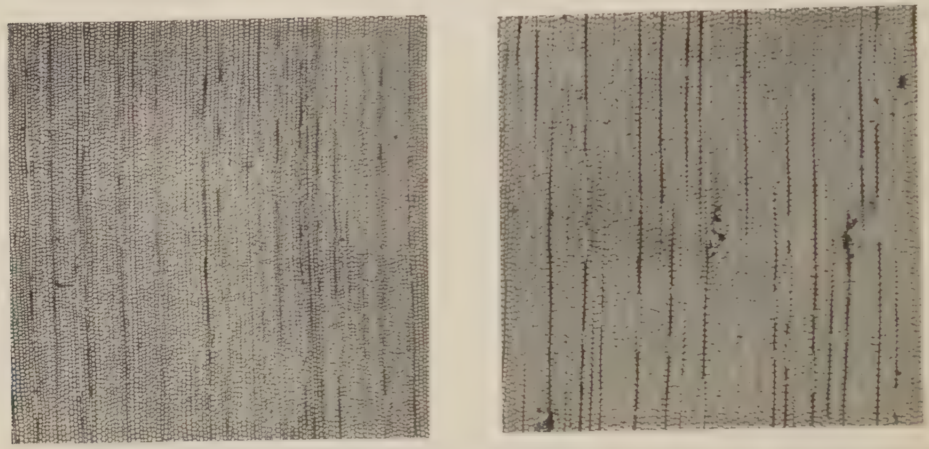
| Section no. | | Percentage transmission of filter | Percentage transmission of section | Cell-space ratio |
|---|---|---|---|---|
| Ash ... | 1 | (10 to 14)<br>(say 12) | 48·5 | 41·5 |
| | 2 | 8 | 45 | 40 |
| | 3 | 17 | 52 | 42·2 |
| | 4 | 14 | 54 | 46·5 |
| | 5 | 10 | 48 | 42·2 |
| | 6 | 10 | 46 | 40 |
| | 7 | 10 | 46 | 40 |
| | 8 | 5 | 45 | 42·1 |
| | 9 | 10 | 50 | 44·5 |
| | 10 | 10 | 48 | 42·2 |
| | | | Average | 42·1 |
| Scots pine | 11 | 2 | 43 | 41·8 |

A set of six matching filters of methylene blue*, passing 2, 5, 8, 10, 14 and 17 per cent of the sodium light, was used.

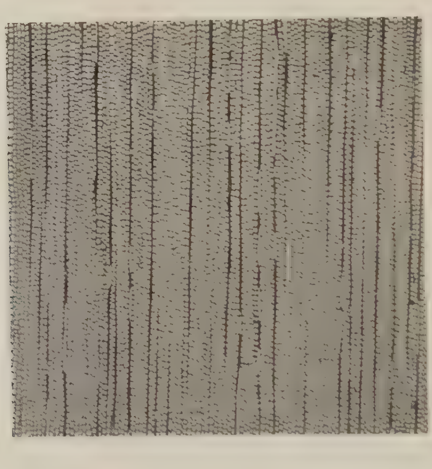
Above about 12-per-cent transmission the choice of matching filter becomes doubtful, and the choice for sections 3 and 4 was difficult for this reason. If these two are omitted on account of matching-errors the average becomes 41·6 per cent and the total scatter 4·5 per cent. This omission is justified since these sections were purposely prepared to cover a range of staining greater than would be expected in practice, where the darkest stain is aimed at and the range, therefore, considerably smaller*.

---

* These filters had been prepared for tests under prism illumination before the sodium vapour lamp had been introduced. Under monochromatic light a set of neutral grey filters would serve equally well and should not be so subject to fading.



Summer wood. Late spring wood.



Early spring wood.

Figure 3. Test sections of *insignis* pine showing areas measured (magnification 17.6).

---

* It should be pointed out that the differences recorded in table 2 for the cell-space ratio are to some extent accounted for by the fact that the sodium-vapour lamp was running from the mains which were subject to sudden changes of load within the Laboratory. Had large-capacity h.-t. accumulators been available a better agreement would have been anticipated. Indeed previous comparison experiments, not recorded here, were carried out with the prism set up with a 12-V. car battery as described above, and proved that a constant source of illumination did give results which were individually more reproducible than was any one reading in table 2. In spite of this advantage the prism method was abandoned in favour of the sodium lamp because, for reasons explained on p. 548 above, the difficulty of matching under prism illumination made for poorer agreement between the values of the cell-space ratio for the separate sections. Moreover, persistent "storms" in the mains are easily observable on the photometer, and readings may be postponed till load-changes are a minimum, whereas the discrepancies under prism illumination, due to colour-changes in the stain, may pass quite unsuspected unless a large number of identical sections are prepared together for comparison. For routine work this is out of the question.

(iii) and (iv). As a test of the agreement between the judgment of match made by several observers, and of the significance of the differences between values obtained on sections having different cell-space ratios, the following results on *insignis* pine, table 3, may be quoted. In this case the values of the lightly stained sections have been omitted as unreliable, though those quoted are, on the whole, lighter than could have been obtained without difficulty.

Table 3. *Insignis* pine. Comparison of cell-space ratio for sections stained to medium depths and matched by observers X, Y and Z

| Section no. | $Q_2/Q_0$ | Transmission of matching filter | | | | Cell-space ratio | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | X | Y | Z | Average | X | Y | Z | Average |
| Summer wood | | | | | | | | | |
| 1 | 46 | 8 | 5 | 8 | 7 | 41 | 43 | 41 | 41·7 |
| 2 | 51 | 11 | 21 | 23 | 18 | 45 | 38 | 31·5 | 38·2 |
| 3 | 47 | 8 | 11 | 8 | 9 | 42 | 40 | 42 | 41·3 |
| 4* | 43 | 8 | 5 | 5 | 6 | 38 | 40 | 40 | 39·3 |
| Averages ... | | | | | | 41·5 | 40·3 | 38·6 | 40·1 |
| Late spring wood | | | | | | | | | |
| 1 | 67 | 5 | 9 | 11 | 8 | 65 | 63 | 62 | 63·3 |
| 2 | 67 | 8 | 7 | 9 | 8 | 64 | 64 | 63 | 63·7 |
| 3 | 65 | 8 | 8 | 8 | 8 | 62 | 62 | 62 | 62·0 |
| Averages ... | | | | | | 63·7 | 63·0 | 62·3 | 63·0 |
| Early spring wood | | | | | | | | | |
| 1 | 70 | 5 | 9 | 11 | 8 | 68 | 67 | 65·5 | 67·5 |
| 2 | 71 | 8 | 7 | 9 | 8 | 68 | 68 | 68 | 68·0 |
| 3 | 68 | 8 | 8 | 8 | 8 | 65 | 65 | 65 | 65·3 |
| Averages ... | | | | | | 67·0 | 66·6 | 66·1 | 66·9 |

A sample of *insignis* pine which had very wide annual rings was chosen. The sections were cut from a single ring-width and mounted as before. The magnification was increased so that three separate areas could be measured within this ring, one in the summer wood, one in the late spring wood, and one in the early spring wood, figure 3.

It will be seen that, except for the summer-wood portion of section 2, which is too lightly stained, the values of the cell-space ratio calculated for the separate observers agree to within 3 per cent. The differences between the values for the three areas chosen are also seen to be significant.

The tests given above show that the method may be relied upon to give consistent and significant results.

* The spring-wood portions of section no. 4 were found to be uneven in stain and were therefore rejected.

## § 6. COMPARATIVE TESTS

(i) The degree of reproducibility of the results having been established, it now became necessary to compare the results obtained by the photometric method with those found by other methods of which the one using Shand's micrometer is the most important. In order to do this, section 1 (table 2) was measured by an experienced observer who made three traverses on the Shand's micrometer. The values obtained were 54, 59 and 50 per cent. This result not only showed the large variations to be expected from this method but also suggested that these values might always be higher than those obtained with the photometer. A previous measurement on Scots pine (no. 11, table 2) gave a similar difference; the photometer gave 41·8 per cent and the micrometer 52·4 per cent (average of 3 observations). Two other observers measured this latter section (no. 11) by the micrometer and obtained similar values. It appeared therefore that the micrometer values were not the result of any personal bias on the part of the individual observer. Tests were therefore undertaken to decide which method gave the more accurate value, and possible causes of the discrepancy were investigated.

There were two main reasons for supposing that the photometer method was not at fault. Firstly, the measurements corresponding to cell-space ratio in the case of the paper grid, § 5 (i), gave true results. Secondly, if the accuracy of the correction were doubted the value of $Q_2/Q_0$ in table 2 would still have to be higher than the true cell-space ratio owing to the transmission of the stained cell walls. The transmissions $Q_2/Q_0$ of the sections as a whole were, however, all lower than the average obtained with the micrometer, nos. 3 and 4 being omitted, and their average (which is 48·2 per cent) was considerably lower.

A low value with the photometer could possibly be accounted for by stain having soaked out of the cell wall into the mounting-medium filling the cell spaces. On this assumption the transmission of the contaminated void spaces (which we
$Q_3$ may call $Q_3$) was calculated with the photometer values of $Q_1$ and $Q_2$ as before,
$A'$ the value $A'$ of the cell-space ratio as measured on the micrometer being assumed to be correct. For section no. 1, table 2, we have

$$Q_2 = 41{\cdot}5;\ Q_1 = 12;\ A' = 54$$

so that $Q_3 = 67$ per cent. This means that one-third of the light falling on the void spaces must be stopped out by the contamination in the voids. Such a large contamination should be easily visible, but on careful comparison of the tint of the voids with that of the clear portion of the slide under the comparison microscopes, no difference could be observed between them. The explanation was therefore untenable. Another possible explanation of the lowest photometer value, namely that the light is scattered in passing through the section, has already been mentioned and is shown to be improbable as a result of the tests given in paragraph (iii) below.

(ii) To decide whether the system of traversing tended inherently to give a wrong result, a large number of circles of different known sizes were drawn on a sheet of paper in as random a fashion as possible, so as to cover the sheet without overlapping. From their radii the area covered by the circles was calculated as 54·5 per cent of the area of the paper. Sixteen traverses were made by ruling fine lines in random directions and measuring intercepts. The total number of circles intercepted was 281, which is of the same order as the number of cells crossed on two annual rings. The final average obtained for 16 traverses was 56 per cent, showing a difference of 1·5 per cent, which is well within the limit of accuracy of the experiment. The addition of four more traverses did not appreciably alter this average value.

(iii) In order to obtain independent measurements of the cell-space ratio of the wood section, two methods were adopted. The first method consisted in projecting an enlarged image of section no. 11, table 2, onto millimetre-squared paper and computing the areas of separate cells from the number of millimetre squares covered by the images. The magnification being necessarily large, a single annual ring was measured by six separate projections of small areas spaced diagonally across the ring*. The results are shown in table 4, which also includes measurements on another wood as a check.

Table 4. Comparison of cell-space ratio by three methods

| Section | Cell-space ratio as measured by | | |
|---|---|---|---|
| | projection | photocell | Shand's method |
| Scots pine | 39 | 42 | 52 |
| Sycamore | 30·7 | 31 | 44 |

It will be seen that the results are in better agreement with those of the photometric method than of Shand's method.

The other measurement of the cell-space ratio was made by fitting the Shand's micrometer on to a projection microscope and making the highly magnified image of the section travel across a white screen in which a single pin-prick had been made. A lamp behind the screen illuminated the pin-hole which served as a reference point. Section 1 of table 2 was measured in this way by Mr Mills, who made three traverses and obtained the average result of 41 per cent, which is in close agreement with that obtained with the photometer. A single traverse by another observer gave 42·1 per cent. This again supports the value obtained with the photometer.

(iv) *Summary of comparative tests.* Table 5 summarizes the average results obtained by means of the various methods tried.

This seems to show conclusively that the error lies in the use of the Shand's micrometer for microscopic observation. The most probable explanation seems to be that, since the eye-piece pointer is dark in tone and has finite, though small,

---

* This process is exceedingly tedious and I am indebted to Mr Mills, of the Section of Wood Structure in the Laboratory, for this painstaking work, as also for the measurements made on the Shand's micrometer.

Table 5. Comparison of results for cell-space ratio by different methods

| Method | Cell-space ratio | | | | |
|---|---|---|---|---|---|
| | Ash (section 1) | Scots pine (section 11) | Syca-more | Paper grid 1 (green) | Paper grid 2 (red) |
| Shand's micrometer | 54 | 52 | 44 | — | — |
| Photometer | 41·5 | 41·8 | 31 | 48·5 | 50·7 |
| Projection with Shand's method | 41 | — | — | — | — |
| Projection on squared paper | — | 39 | 30·7 | — | — |
| Travelling microscope | — | — | — | 49·6 | 50·9 |

width, there is a tendency to make this pointer overlap on the *dark* side of the dividing line between cell space and cell wall. This in the extreme case would introduce an error equal to the pointer-width in each cell measured, and this over a large number of small cells will amount to an appreciable figure, giving an excess value to the cell spaces.

## § 7. CONCLUSIONS

The apparatus described above has been used successfully in this Laboratory. If we leave out of account the time required for the sodium light and the amplifier to settle down after switching-on, the time required for the actual measurements on 12 sections is approximately as follows:

| | |
|---|---|
| Selection of matching filters | 60 min. |
| Measurement of $Q_2/Q_0$ and $Q_1/Q_0$ | 40 min. |
| Total | 1 hr. 40 min. |

The time required to measure a section by taking three traverses with the Shand's micrometer is about 20 minutes, so that the time for the twelve sections is 4 hours. The time needed for mounting the sections and computing the results is the same by either method.

## § 8. ACKNOWLEDGMENTS

My thanks are due to members of the staff of the Section of Wood Structure in this Laboratory for their valuable cooperation in this work, and to the Director of the Forest Products Research Laboratory, Mr W. A. Robertson, for permission to publish the paper.

## APPENDIX: AMPLIFIER USED IN THE PHOTOELECTRIC MEASUREMENT OF CELL-SPACE RATIO

The photoelectric current from the photocell is amplified and measured by a null method which is a modification of that described by Campbell and Ritchie*. The circuit is shown in figure 4.

---

* *Photoelectric Cells*, p. 141 (Pitman, 1930).

The bridge, consisting of valves 1 and 2 (chosen with steep characteristics) together with resistances $R_1$ and $R_2$, can be balanced by adjusting $R_{1A}$, with $R_{1B}$ as fine adjustment, together with the potentiometer $P_3$, which alters the ratio of the filament currents in the two valves. If the cell $P$ is dark and the potentiometers $P_{2A}$ and $P_{2B}$ are at zero, each valve has the same grid-bias through $E_g$ 1. Under these conditions the bridge is taken through a series of balances until a small change in the main filament current, made by short-circuiting the resistance $r_1$, produces no change in zero of the galvanometer. If the cell is now illuminated, the bias on $V_1$ becomes less negative owing to conduction through $P$, but may be restored to its former value by tapping off the necessary extra potential from $E_g$ 2 and thus rebalancing the bridge. This extra potential is measured and is proportional to the intensity of the light falling on $P_1$, no calibration of the valves being necessary.

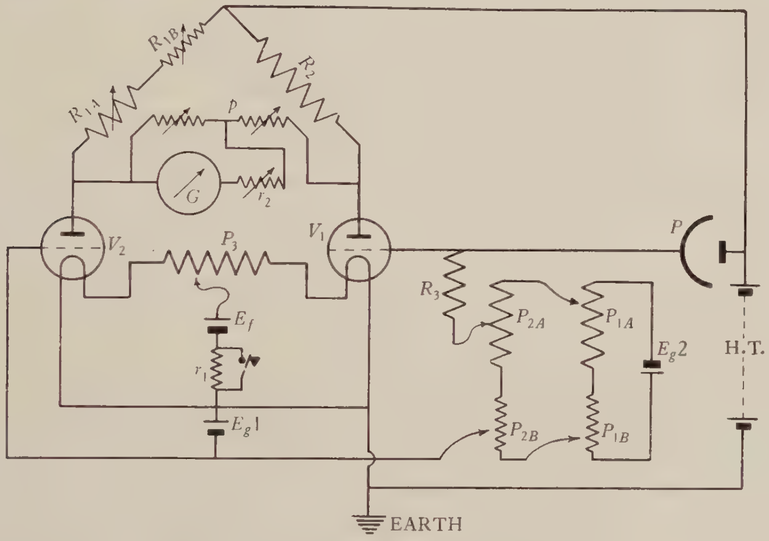


Figure 4. Photoelectric amplifier circuit. $P$, caesium coated vacuum cell giving saturation below 120 V.; $V_1$ and $V_2$, Mullard. P.M. 1.H.L. 2–150 V.; $r_1$, 1 Ω.; $R_{1A}$ and $P_{1A}$, 20,000 Ω.; $R_2$, 15,000 Ω.; $R_{1B}$ and $P_{1B}$, 400 Ω.; $R_3$, 2 to 8 MΩ. according to sensitivity required; $P_{2A}$, 100 steps of 200 Ω.; $P_{2B}$, 200 Ω. (dial marked in 20-Ω. steps); $P_3$, 4 Ω.; $E_g$ 1, 1·5-V. dry battery; $E_g$ 2 and $E_f$, 2-V. accumulators; H.T., 120 V.

$P_{2A}$ is a 100-stud potentiometer having steps of 200 Ω., specially made by Messrs Gambrell Bros. $P_{2B}$ is a slide-wire potentiometer with a total resistance of 200 Ω. $P_{2A}$, with $P_{2B}$ for steps of less than 1 per cent, records directly the intensity of any light as a percentage of a brighter one* and is used as follows: After the dark balance has been obtained $P_{2A}$ is put at 100 per cent and $P_{2B}$ at zero, the cell is fully illuminated through the clear portion of the microscope slide, and the balance of the bridge is restored by $P_{1A}$, with $P_{1B}$ as fine adjustment, giving $Q_0 = 100$ per cent. $P_1$ is now kept constant, and any other, smaller, intensity

---

* Strictly speaking part of $P_{2A}$ is shunted by the current through the grid-leak $R_3$ and the photocell $P$, but as this current is extremely small even at full illumination it may be disregarded, and equal increments of $P_{2A}$ may be taken as equal increments of the potential on the grid of the valve $V_1$.

$Q_1$ or $Q_2$ may be compared with the first by balancing with $P_{2A}$ only, and $P_{2B}$ for steps of less than 1 per cent. This device saves much calculation and gives directly the value of the ratio required. Moreover, changes in the intensity of the source are at once noticed, since the setting of $P_1$ for 100-per-cent illumination is independent, in relation to the dark zero, of changes in the rest of the circuit.

Frequent checking of the dark setting is necessary to correct a tendency to slow drift of zero. This drift is in part caused by any local temperature-changes and may be reduced by lagging the casing of the amplifier.

The sensitivity of the galvanometer may be adjusted to suit any maximum intensity of light by means of a potentiometer and series resistance, $p$ and $r_2$, figure 4, and is set to give the most rapid dead-beat swing with a deflection of about 5 mm. to 10 mm. for a 1-per-cent change in light*. The amplifier is enclosed in a light-tight earthed metal case. As an extra precaution against stray light the photocell is also cased in black paper which is in contact with the cell only at the base.

## DISCUSSION

Lord RAYLEIGH said that in measuring the area of the holes in wire gauze he had obtained by far the most convincing results by means of a photoelectric cell. Photometric measurements used to be regarded as the most difficult of all, but modern equipment had brought such measurements within the domain of electrical measurement, and electrical measurements are in general easier to carry out than any other.

Mr J. E. CALTHROP enquired whether the method was applicable to small particles, such as lycopodium powder with a diameter of about 0·03 mm., or would diffraction vitiate the results?

Mr J. GUILD suggested that the transmitted light might be collected in an integrating sphere in order that errors due to diffraction might be eliminated.

Mr T. SMITH. The author has established a method for the examination of wood which he finds to be at once more rapid and more accurate than those used previously. The only surprising feature is that with so large a number of small cells the author has found no sign of diffraction. The answer to the question concerning the mutual effect of the cells on diffraction is that each aperture diffracts as if the others were absent, but the size and arrangement of the cells as a whole affect the distribution of the diffracted light in space.

Dr E. G. RICHARDSON. I should like to ask whether the apparatus would be suitable for a ready determination of the suitability of a given specimen of wood to support a load under unfavourable conditions, such as wet surroundings? I am thinking particularly of the testing of wood for use as pit props. With regard to the difficulty of instituting a comparison test of the method, to which the author refers, he might have examined the light passing through a suspension of sand in water vigorously stirred in a glass tank. It has been shown both theoretically and by pipette sampling that with complete turbulence the sand is distributed in the water exponentially with respect to height*. Such a suspension is, of course, the obverse of the author's microsections, but I have found that the light transmitted through such a system at different heights, with reference to clear water as standard, *is* proportional to the theoretical distribution as long as the concentration is not too high; and this would seem to indicate if the two experiments are comparable that the author may allay his fears with regard to scattered light upsetting the results. The trouble with regard to change of intensity of source during an experiment may be obviated by using two well-matched cells in a sort of photoëlectric Wheatstone's Bridge.

---

* In cases where a brighter light-source can be used, the galvanometer may with advantage be replaced by a dial-pattern centre-zero microammeter (e.g. Gambrell's) which greatly increases the speed at which measurements can be made.

AUTHOR'S reply. If the transmitted light were measured by means of an integrating sphere, light scattered from the translucent cell wall would be picked up as well as that from the holes and would have to be allowed for in the choice of matching filter. This could not be done unless the filter were of material identical, except for the holes, with that of the section, which is hardly practicable. In the present form of the apparatus this cell-wall scattering is allowed for, since the matching filter is chosen in respect of the light transmitted normally, and whether the rest of the light is scattered or absorbed is of no importance.

I am comforted by Dr Richardson's experiments showing that the scattered light does not upset his results. I think the reason is that the area of the suspension (or, in my case, of the section) which is illuminated is considerably larger than the area under test whose image falls on the aperture of the photocell. Thus as much light will be scattered on to the aperture from the surrounding suspension as away from it by the area under examination. Provided that the scattered light is roughly the same from all parts of the suspension, this will give an automatic compensation for scattering up to considerable angles from the normal.

Owing to the small amount of use to which this apparatus has so far been put, it is premature to say how far it would assist Dr Richardson in his problems of pit props, but from other work on pit props which has been done in this Laboratory† it would seem that the cell-space ratio within the timber is not by any means the only factor determining its suitability for pit props.

---

* Hurst, *Proc. R.S.* **124**, 196 (1929).

† "Home-grown pit props," *F.P.R.L. Bulletin* No. 9 (1931).

519.281

# THE MID-COURSE METHOD OF FITTING A PARABOLIC FORMULA OF ANY ORDER TO A SET OF OBSERVATIONS

By T. SMITH, M.A., F.Inst.P., F.R.S.

*Received March* 16, 1934. *Read May* 18, 1934.

*ABSTRACT.* A meaning for the closeness of fit of a formula which is considered widely applicable in physics is suggested, and a method of deriving the most closely fitting parabolic formula of a given order is described. Geometrically the formula represents the mid course between the most outlying points.

## § 1. PRINCIPLES OF THE METHOD

It is no rare occurrence in physics for the problem of fitting a formula of the type

$$y = a + bx + cx^2 + \dots + kx^{n-1},$$

where corresponding values of $x$ and $y$ are given, to arise, and a number of papers have been written on the procedure to be followed. These show that there is considerable diversity of opinion on the way in which the problem should be solved. Some methods, for instance graphical methods, are advocated because of the simplicity of the numerical work involved, and others, for example the method of least squares, because they are claimed to lead to the most accurate results. But even when accuracy is admittedly of outstanding importance, different methods, leading to somewhat different results, are advocated by different writers. I suggest that the reason for this disagreement is to be found in the absence of any clear statement of what is to be understood by accuracy. There is no single criterion of accuracy or closeness of fit which is of general applicability. It will, I think, be generally conceded that there are problems in which a stationary value of $\Sigma\,(\Delta y)^2$ is the criterion which ought to be applied. In others this sum may perhaps requre to be replaced by $\Sigma\,\{A\,(\Delta x)^2 + 2B\Delta x\Delta y + C\,(\Delta y)^2\}$, where $A$, $B$, $C$ are constants which depend upon factors special to the particular problem and upon the scales on which $x$ and $y$ are measured. But in many physical problems no variant of the method of least squares is of special interest as a means of constructing a formula of the type we are considering. The more usual requirement is that for any value of $x$ in a given range the formula should give $y$ correct within certain fixed limits. If $y$ is required to be correct to three significant figures, to point out that the formula is almost everywhere correct to four or five figures is no answer to the objection that in one or two places it is wrong in the second figure. Since this is the type of error-compensation to which the least-squares method may lead, we clearly ought only to employ this method when an examination of the problem shows that a solution of this kind is required. In physics the value of a formula depends much more frequently on the size of the maximum error to which its use can lead.

The relative values of empirical formulae of the polynomial type may be measured roughly by the inverse of the number of terms they contain. For this reason we almost always employ a shorter expression than the one which would represent the observations exactly. The formula then in general yields values of $y$ differing by smaller or larger amounts from the actual experimental results. The problem to be considered here is the determination of the coefficients $a$, $b$, $c$, ... so that none of the larger discrepancies exceed a minimum value; the smaller discrepancies will not trouble us. In other words, we completely neglect all the observations that fall in mean positions and determine our coefficients from the outlying observations alone. If the extreme errors prove to be too great the order of the polynomial must be increased. The outlying observations can usually be selected without difficulty by plotting $y$ against $x$, but it may sometimes be necessary to plot $y - \phi(x)$ against $x$, where $\phi(x)$ is an approximate expression for $y$. The number of outlying points should be one more than the number of terms in the polynomial, and therefore two more than the order of the polynomial. As $x$ increases from its lowest to its highest value they must be chosen to lie alternately on opposite sides of the mean locus of points.

Assume for the moment that we can find values for the $n$ coefficients so that

$$f(x) \equiv a + bx + cx^2 + \dots + kx^{n-1}$$ *f*

yields values of $y$ for the outstanding observations which are all in error by the same amount but alternately of opposite signs. Thus if $x_1, x_2, x_3, \dots x_{n+1}$ are the values of the independent variable for the selected observations, arranged as an ascending or as a descending sequence, we have

$$y_1 - f(x_1) = f(x_2) - y_2 = y_3 - f(x_3) = f(x_4) - y_4 = \dots = \epsilon.$$ $\epsilon$

Then $f(x)$ is in error by $-\epsilon$ at $x_1, x_3, x_5, \dots$, and by $+\epsilon$ at $x_2, x_4, x_6, \dots$, and by amounts not exceeding these limits at all the other observed points.

By an argument which has been given before* in a somewhat different problem but which may be worth repeating here, it can be shown that $f(x)$ is the polynomial of order $n - 1$ which fits the observations best. If possible suppose that $g(x)$ is a *g* polynomial of order not greater than $n - 1$ which represents $y$ with errors not exceeding the narrower limits $\pm |\eta|$. Then if $\epsilon$ is positive $\eta$

$$f(x) - g(x) \leqslant -\epsilon + |\eta| < 0 \text{ for } x = x_1, x_3, x_5, \dots$$

and

$$f(x) - g(x) \geqslant \epsilon - |\eta| > 0 \text{ for } x = x_2, x_4, x_6, \dots.$$

If $\epsilon$ is negative the same results hold with the interchange of the sets $x_1, x_3, x_5, \dots$ and $x_2, x_4, x_6, \dots$. In either event it follows that the polynomial $f(x) - g(x)$, which at most is of order $n - 1$, vanishes for at least $n$ different values of $x$ in the interval $x_1$ to $x_{n+1}$. But this is only possible if $f(x)$ and $g(x)$ are identically equal, and this is inconsistent with the supposition that $|\eta|$ is less than $|\epsilon|$. The polynomial $f(x)$ is therefore the one which yields the least extreme errors in $y$, and is therefore the expression desired.

---

* *Proc. Phys. Soc.* **33**, 141 (1919–20).

The actual construction of the polynomial is very simple. Let

$F$

$$F(y_1, y_2, y_3, y_4, \ldots), \text{ standing for}$$

$$\sum_{r=1}^{n+1} \frac{y_r (x-x_1)(x-x_2)\ldots(x-x_{r-1})(x-x_{r+1})\ldots(x-x_{n+1})}{(x_r-x_1)(x_r-x_2)\ldots(x_r-x_{r-1})(x_r-x_{r+1})\ldots(x_r-x_{n+1})},$$

be the Lagrangian polynomial of order $n$ which passes through the $n - 1$ given points. Then, if $x_1, x_2, x_3, \ldots$ are regarded as fixed, $F(1, -1, 1, -1, \ldots)$ is a polynomial of order $n$ which yields the values $1, -1, 1, -1, \ldots$ at $x_1, x_2, x_3, x_4, \ldots$. Let the coefficient of $x^n$ in $F(y_1, y_2, y_3, y_4, \ldots)$ be $\epsilon$ times the coefficient of $x^n$ in $F(1, -1, 1, -1, \ldots)$. Then

$$f(x) = F(y_1, y_2, y_3, \ldots) - \epsilon F(1, -1, 1, -1, \ldots)$$

is a polynomial, of order $n - 1$ at most, which assumes the values

$$y_1 - \epsilon, y_3 - \epsilon, y_5 - \epsilon, \ldots \text{ at } x = x_1, x_3, x_5, \ldots$$

and the values

$$y_2 + \epsilon, y_4 + \epsilon, y_6 + \epsilon, \ldots \text{ at } x = x_2, x_4, x_6, \ldots.$$

That is to say $f(x)$ is the expression required, and $\epsilon$ is the magnitude of the extreme error. It may be observed that though $F(y_1, y_2, y_3, \ldots)$ is of higher order than $f(x)$ it is liable to yield errors amounting to about $2\epsilon$, i.e. twice as great as those possible with $f(x)$.

$(r)_p$ If it is preferred we can write down explicit formulae for $\epsilon, a, b, c, \ldots$. Let $(r)_p$ denote the sum of the homogeneous products of the $x$'s taken $p$ at a time, $x_r$ being excluded. Denote

$Y_p$

$$\sum_{r=1}^{n+1} \frac{y_r (r)_p}{(x_1 - x_r)(x_2 - x_r)\ldots(x_{n+1} - x_r)} \text{ by } Y_p$$

$X_p$ and

$$\sum_{r=1}^{n+1} \frac{(-1)^r (r)_p}{(x_1 - x_r)(x_2 - x_r)\ldots(x_{n+1} - x_r)} \text{ by } X_p,$$

where the denominators represent the continued product of the excess of all the other $x$'s over $x_r$. Then if the sign of $\epsilon$ is unchanged from the preceding expressions the formulae required are

$$\begin{aligned} -\epsilon &= Y_0/X_0, \\ a &= Y_n + \epsilon X_n, \\ -b &= Y_{n-1} + \epsilon X_{n-1}, \\ c &= Y_{n-2} + \epsilon X_{n-2}, \\ -d &= Y_{n-3} + \epsilon X_{n-3}, \end{aligned}$$

and so on. The first of these equations may be used to determine whether the $n - 1$ observations have been chosen correctly. The choice is correct if the substitution of any other observations for one or more of those selected yields a smaller absolute value of $\epsilon$.

If formulae cast into a particular mould are desired it is best to start from these equations. As an example the divided-difference formulae for fitting an ordinary parabola are

$$c = \frac{\dfrac{y_1 - y_3}{x_1 - x_3} - \dfrac{y_2 - y_4}{x_2 - x_4}}{x_1 - x_2 + x_3 - x_4},$$

$$b = -\frac{\dfrac{y_1 - y_3}{x_1 - x_3}(x_2 + x_4) - \dfrac{y_2 - y_4}{x_2 - x_4}(x_1 + x_3)}{x_1 - x_2 + x_3 - x_4},$$

$$a + \epsilon = \eta_{13} + cx_1x_3, \quad a - \epsilon = \eta_{24} + cx_2x_4,$$

where $\eta_{13}$, $\eta_{24}$ are the ordinates of the points in which the chords 1, 3 and 2, 4 meet the $y$ axis, so that

$$\eta_{13} = \frac{x_1y_3 - x_3y_1}{x_1 - x_3} = \frac{y_1 + y_3}{2} - \frac{y_1 - y_3}{x_1 - x_3} \cdot \frac{x_1 + x_3}{2},$$

$$\eta_{24} = \frac{x_2y_4 - x_4y_2}{x_2 - x_4} = \frac{y_2 + y_4}{2} - \frac{y_2 - y_4}{x_2 - x_4} \cdot \frac{x_2 + x_4}{2}.$$

For a cubic the simplest formula is perhaps

$$f(x) + \epsilon = y_1 \frac{(x - x_3)(x - x_5)}{(x_1 - x_3)(x_1 - x_5)} + y_3 \frac{(x - x_1)(x - x_5)}{(x_3 - x_1)(x_3 - x_5)} + y_5 \frac{(x - x_1)(x - x_3)}{(x_5 - x_1)(x_5 - x_3)}$$
$$+ \frac{2(x - x_1)(x - x_3)(x - x_5)}{x_1^2 - x_2^2 + x_3^2 - x_4^2 + x_5^2 - (x_1 - x_2 + x_3 - x_4 + x_5)^2} \left[ \frac{(y_1 - y_3)(x_2 - x_3 + x_4 - x_5)}{(x_1 - x_3)(x_1 - x_5)} \right.$$
$$\left. - \frac{y_2 - y_4}{x_2 - x_4} + \frac{(y_3 - y_5)(x_1 - x_2 + x_3 - x_4)}{(x_5 - x_1)(x_5 - x_3)} \right].$$

The full equation for $\epsilon$ is

$$\{x_1^2 - x_2^2 + x_3^2 - x_4^2 + x_5^2 - (x_1 - x_2 + x_3 - x_4 + x_5)^2\} \epsilon$$
$$= \frac{y_1(x_2 - x_3)(x_2 - x_5)(x_3 - x_4)(x_4 - x_5)}{(x_1 - x_3)(x_1 - x_5)} - \frac{y_2(x_1 - x_4)(x_3 - x_4)(x_4 - x_5)}{(x_2 - x_4)}$$
$$+ \frac{y_3(x_1 - x_2)(x_2 - x_5)(x_1 - x_4)(x_4 - x_5)}{(x_1 - x_3)(x_3 - x_5)} - \frac{y_4(x_1 - x_2)(x_2 - x_3)(x_2 - x_5)}{(x_2 - x_4)}$$
$$+ \frac{y_5(x_1 - x_2)(x_1 - x_4)(x_2 - x_3)(x_3 - x_4)}{(x_1 - x_5)(x_3 - x_5)},$$

but usually it will be more easily obtainable by substituting $x_2$ or $x_4$ for $x$ after the numerical expression for $f(x) + \epsilon$ has been found, and using

$$\epsilon = f(x_2) - y_2 = f(x_4) - y_4.$$

As the order of the curve increases it becomes less profitable to construct special algebraic formulae. The least amount of computation will probably then follow from the restriction of the homogeneous sums to the case when all the points are included. This is accomplished by writing the polynomial formally as the product

$$(x - x_1)(x - x_2) \dots (x - x_{n+1})(x^{-2}A_1 + x^{-3}A_2 + x^{-4}A_3 + \dots),$$

it being understood that in the final product negative powers of $x$ are to be omitted. The $A$'s (including two additional quantities $A_0$ and $A_{-1}$) are then found from

$$A_p = \Sigma \frac{y_r x_r^p}{\pi_r} + \epsilon \Sigma \frac{(-1)^r x_r^p}{\pi_r},$$

where $$\pi_r = (x_r - x_1)(x_r - x_2) \dots (x_r - x_{n+1})$$

and $\epsilon$ is given by $A_0 = 0$. The constant term can also be evaluated directly as $x_1 x_2 \dots x_{n+1} A_{-1}$. The calculations can be checked in the course of the work from the identities

$$\Sigma \frac{x_r^p}{\pi_r} = 0 \qquad (p = 0, 1, 2, \dots n-1),$$

$$\Sigma \frac{1}{x_r \pi_r} = \frac{(-1)^n}{x_1 x_2 \dots x_{n+1}}, \quad \Sigma \frac{x_r^n}{\pi_r} = 1,$$

which are also useful in reducing the labour involved in constructing special formulae such as those already given for the cubic case.

## § 2. NUMERICAL EXAMPLES

It is of interest to see how results obtained by this mid-course method compare with those given by other methods. We take first of all a simple example given by Merriman. The abscissae and ordinates of four points on a straight line were measured with equal precision, giving

$$y = 0·5,\ 0·8,\ 1·0, \text{ and } 1·2,$$

$$x = 0·4,\ 0·6,\ 0·8, \text{ and } 0·9.$$

By considering errors in $x$ and $y$ simultaneously Merriman reaches the least-squares solution

$$y = 1·348x - 0·035.$$

From a graph it is easily seen that the first point lies closer to any reasonable mean line than the other three. The slope of the mid-course line is thus determined from the second and fourth points and is $(1·2 - 0·8)/(0·9 - 0·6)$ or $4/3$, and it at once follows that the line given by the method of this paper is

$$y = (40x - 1)/30.$$

For the given values of $x$ the errors in $y$ by the two methods are

| | | | | |
|---|---|---|---|---|
| Least squares | 0·0042, | − 0·0262, | 0·0434, | − 0·0218. |
| Mid course | 0·0000, | − 0·0333, | 0·0333, | − 0·0333. |

If the points and these lines are plotted it is not possible from the graph to say that one is a better fit than the other. The squares of the errors sum to 0·00306 and 0·00333 respectively, so that even from the least-squares point of view the solution we have found with such little trouble cannot be substantially improved.

If we did not know that the points were supposed to be on a straight line we might consider whether a parabola would give a better fit. The reduced formulae give

$$c = \tfrac{10}{36},\ b = \tfrac{11}{12},\ a \pm \epsilon = c \times 0·32,\ a \mp \epsilon = c \times 0·54$$

or $\epsilon = 0{\cdot}0305$. The reduction in the error is therefore trivial, and there is no justification for preferring a parabola to a straight line.

The next example from Merriman has been utilized by Awbery. The temperatures encountered at different depths in boring a deep Artesian well were found to be

| Depth $x$ (m.) | 28 | 66 | 173 | 248 | 298 | 400 | 505 | 548 |
|---|---|---|---|---|---|---|---|---|
| Temperature $y$ (°C.) | 11·71 | 12·90 | 16·40 | 20·00 | 22·20 | 23·75 | 26·45 | 27·70 |

From a graph the first, third, fifth and sixth points were selected. The mid-course solution found for an ordinary parabolic formula is

$$y = 9{\cdot}694 + 0{\cdot}047395x - 0{\cdot}000026206x^2.$$

Awbery obtained by the zero-sum method

$$y = 9{\cdot}745 + 0{\cdot}049495x - 0{\cdot}0000322x^2,$$

and I find for the least-squares solution

$$y = 10{\cdot}214 + 0{\cdot}044227x - 0{\cdot}000023115x^2.$$

(Merriman's solution, $y = 10{\cdot}6 + 0{\cdot}0415x - 0{\cdot}0000193x^2$, is subject to the condition $y = 10{\cdot}6$ when $x = 0$.)

The errors due to the use of these formulae are given in table 1.

Table 1

| $x$ | Mid course | Zero sum | Least squares |
|---|---|---|---|
| 28 | − 0·71 | − 0·60 | − 0·28 |
| 66 | − 0·19 | − 0·03 | + 0·13 |
| 173 | + 0·71 | + 0·94 | + 0·77 |
| 248 | − 0·16 | + 0·04 | − 0·24 |
| 298 | − 0·71 | − 0·56 | − 0·86 |
| 400 | + 0·71 | + 0·64 | + 0·46 |
| 505 | + 0·50 | + 0·08 | + 0·20 |
| 548 | + 0·10 | − 0·50 | − 0·19 |

From these figures it is apparent that the zero-sum curve, which almost passes through the second, fourth and seventh points, is nearer the least-squares solution than the mid-course solution, though the parabolas for the latter two agree much more closely in the position of the axis and the length of the latus rectum than either of them do with the zero-sum solution.

As a final example from Merriman we take one in which the accuracy of fit by a parabolic formula corresponds more closely to accurate physical measurements. The data give the velocities $v$ of the water of the Mississippi at different depths $d$: *v, d*
they, and the outstanding errors with the mid-course and least-squares solutions, are given in table 2.

Under these conditions there is little to choose between the solutions, and the chief advantage of the mid-course method is the relatively small amount of labour involved in reaching an acceptable solution.

The problem we have been considering is essentially that known to the statis-

Table 2

| $d$ | $v$ | Mid course | Least squares |
|---|---|---|---|
| 0·0 | 3·1950 | + 0·0027 | − 0·0001 |
| 0·1 | 3·2299 | + 0·0036 | − 0·0018 |
| 0·2 | 3·2532 | + 0·0009 | − 0·0002 |
| 0·3 | 3·2611 | − 0·0016 | − 0·0021 |
| 0·4 | 3·2516 | − 0·0020 | − 0·0019 |
| 0·5 | 3·2282 | − 0·0036 | − 0·0031 |
| 0·6 | 3·1807 | + 0·0036 | − 0·0044 |
| 0·7 | 3·1266 | + 0·0022 | − 0·0033 |
| 0·8 | 3·0594 | − 0·0012 | 0·0000 |
| 0·9 | 2·9759 | − 0·0036 | − 0·0024 |

tician as "smoothing," with the restriction that the smooth curve is to be represented by a simple algebraic formula. It becomes of interest to consider how the results obtained by the method of this paper compare with those reached by approved

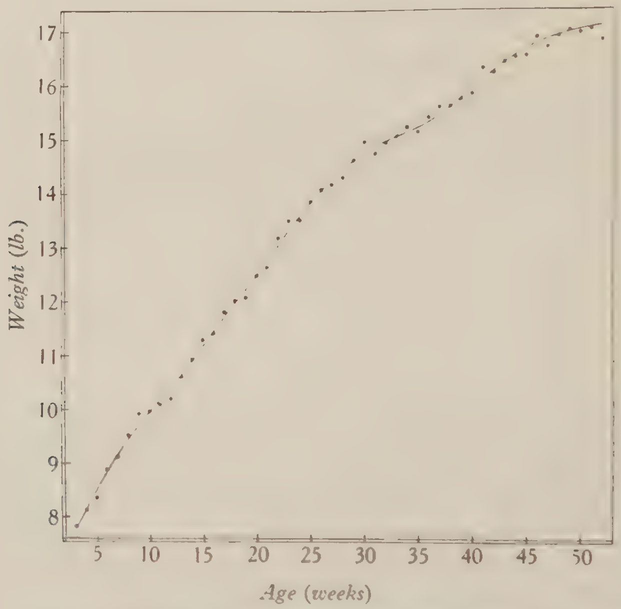


Figure 1. Weight of babies plotted against age, with Rhodes's smoothed values shown as a continuous curve.

methods of smoothing. For this purpose I have taken two of the examples given by Rhodes* in his tract on smoothing. The first of these represents the weight of babies according to age for their first year. The observations and Rhodes's smoothing are shown in figure 1. Actual values are not however quoted for the first seven and the last six points plotted, so that I have been forced to consider only the central three-quarters of the plot. The points shown ringed in figure 2 were chosen by eye as likely to be the most outstanding, and as they were four in

* *Tracts for Computers*, No. 6 (Cambridge University Press).

number an ordinary parabola was fitted. The curve is seen to be a good general fit and it happens to fit equally well over the whole range of the points shown in figure 1. The question now arises whether the departures of the points from the curve ought to be treated as accidental or not. If the measurements referred to a normal physical problem I should consider any attempt to improve upon this result as unjustified. We can do a little to bring the curve nearer the majority of the points by adding the first point to the other four and fitting a cubic, but the improvement is so small that it is doubtful whether it is worth making. If the undulations in the curve of figure 1 have a meaning, figure 2 illustrates a stage beyond which the mid-course method as it has been developed here is inapplicable.

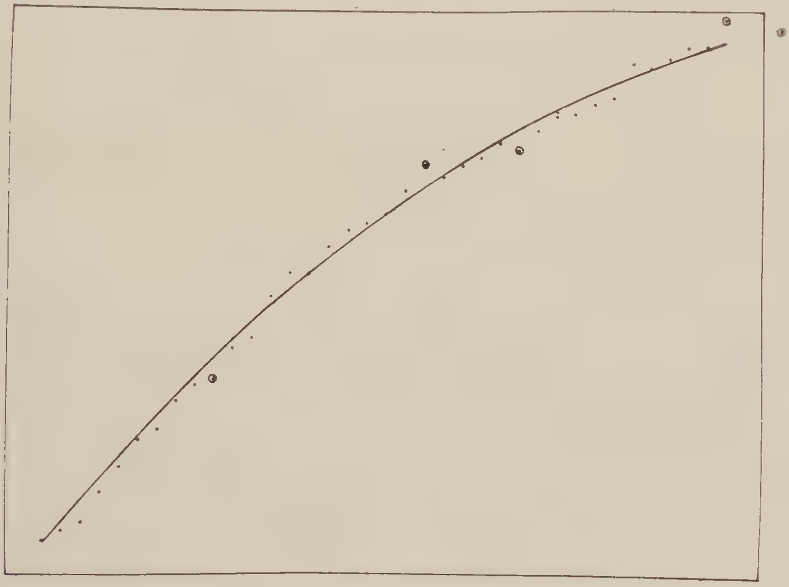


Figure 2. The mid-course parabola fitted to the points of figure 1..

Figure 3 is an example where the scatter of the points is considerable. The points represent variation of mortality with age, and their distribution suggests a cubic curve to the eye. It is easy to pick out five points likely to represent extreme departures from the curve: they are distinguished in the figure by large circles. The cubic curve based on these points extends from near the top left-hand corner to the bottom right-hand corner. Though the curve agrees very fairly with the run of the points, it is apparent that considerable improvement is possible in the right half of the diagram but that the points on the left are too scattered to leave correspondingly small residuals. By adding the two points marked with smaller circles we push the curve further away from these two points, while bringing it nearer to the five points originally selected. The new quintic curve is shown by a second full line, and, with the exception of the first third, represents the data very satisfactorily. The smoothing may be extended without difficulty—say by means of a cubic curve

—into the remaining part of the diagram with extreme errors somewhat less than those obtained with the original cubic. No substantial improvement on this result is likely to be obtainable, and we may infer from this and the preceding example

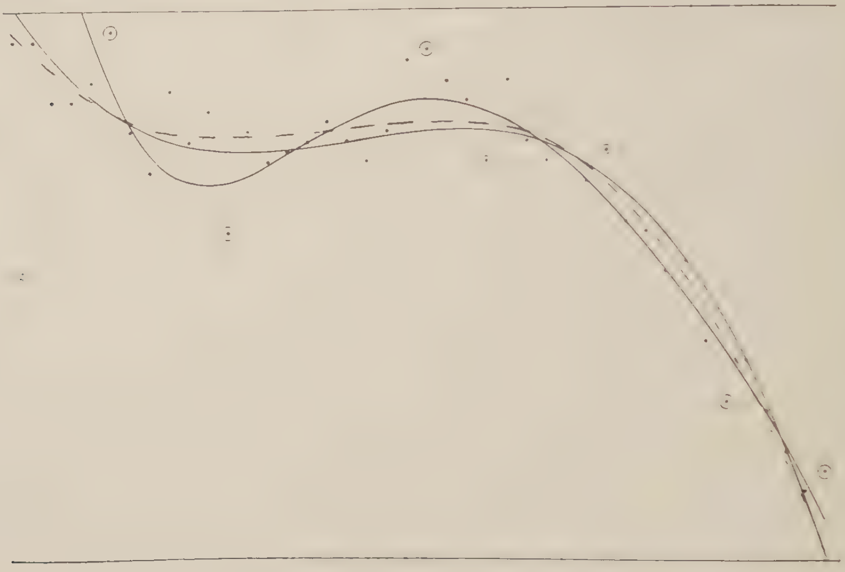


Figure 3. The full lines represent the mid-course cubic and quintic; the broken line represents the least-squares cubic.

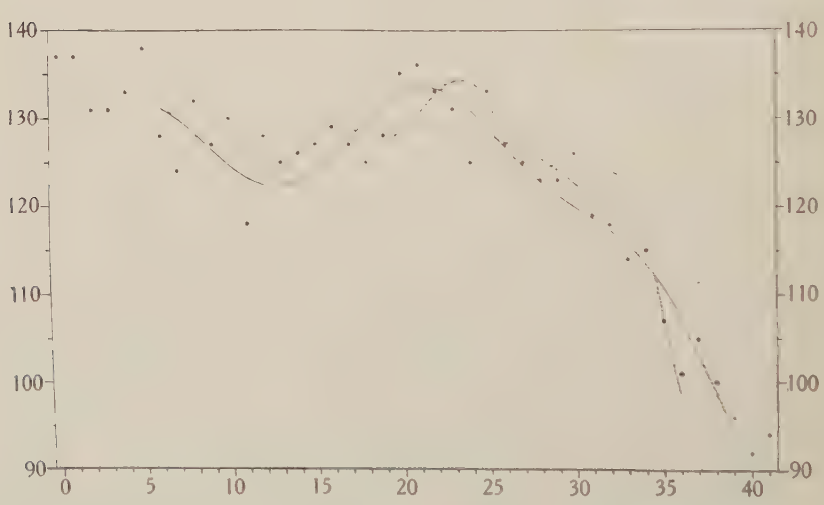


Figure 4. Smoothing of mortality curve. ······ cubic; –—— osculatory method A; ------ osculatory method B; ••••• original data.

that the mid-course method is not unlikely to be applicable in some statistical problems where it may well be held that the method of least squares has clearer justification than in some branches of physics.

The broken curve of figure 3 represents approximately the cubic curve obtained by the method of least squares. This curve is shown accurately in figure 4, which is taken from Rhodes's pamphlet*, where the results of smoothing in the central region by two other methods are shown. From one of the examples it is clear that smoothing can be carried out in ways which lead to mere misrepresentation of the significant features of the data. This is not possible with the mid-course method if it is used intelligently. The other curve is free from faults of this kind, but it certainly does not represent the observations better than the curves obtained here, though its derivation involves much more labour.

## § 3. ACKNOWLEDGMENT

I have to thank the Director of the Department of Applied Statistics, University College, London, for permission to reproduce figures 1 and 4 from Rhodes's tract, to which reference has already been made.

## DISCUSSION

Lord Rayleigh said that the paper perhaps hardly took sufficient account of the difference between an error and a mistake. Suppose that fifty observations had the value $A + dA$ and one the value $A - dA$. According to the author the best value would be $A$, but this conclusion would hardly appeal to a betting man, and still less would it do so if the fifty-one observations were made by fifty-one different observers. In the case of heterochromatic photometry some observers were very confident while others were positively tortured by doubt when making a setting. Probably there was that element of doubt, to a greater or less extent, in all observations. The main use of methods of reducing observations by least squares and the like was to enable a man to state his results succinctly without faking them. It might be doubted whether such methods had ever produced real knowledge in the shape of a discovery unless perhaps in proving the reality of stellar parallax. The speaker had been taught in his youth that the theory of errors of observation was a thing to be learned and then forgotten. But at all events such methods as that of the author would save weary hours of arithmetic.

Mr J. H. Awbery. Mr Smith has put forward a criterion for accuracy of fit which has much to recommend it, and merits careful consideration. There will doubtless be many cases—the derivation of a formula to represent the observed (mass, luminosity) relation in stars is perhaps one—where it is the best criterion, but I do not think it will often be suitable in pure physics. Take for example the extreme case of a formula of degree zero in the independent variable in which we have a number of estimates of a single quantity. Mr Smith's method would take as a representative value the mid-point between the greatest and least of the estimates. If these are measurements of a physical quantity, my objections to this are (1) that a single further repetition which gives a bad result can alter the representative value, and (2) that no account is taken of the possibility of mistakes, as distinct from errors. That Mr Smith sees the matter differently is clear, when he speaks of the case where a formula is almost everywhere accurate to four or five figures, but is wrong in the second figure at a few points. I agree with him, that the formula is condemned, if the observations are indisputable; but if they were measurements of a pair of physical variables, I should draw the conclusion that observations ought to be repeated at the points where the large discrepancies occur.

---

* In the original diagram about one-quarter of the points are incorrectly plotted. These errors have been corrected in this diagram.

It is clear that the selection of the best criterion depends on the nature of the problem; it is probably also in some degree a matter of taste, depending on one's psychological make-up. However that may be, we must be grateful to Mr Smith for his very skilful formulation of the method to be followed in deriving the formula, on those occasions when it has been decided that his criterion is to be adopted.

Mr T. W. Adams. Would Mr Smith say what he regards as the limitations of his method of computation? Does he regard the method as applicable to physiological and biological data, i.e. to vital statistics?

Mr J. Guild. I think that the difficulty which some participants in the discussion find in accepting the mid-point criterion put forward by Mr Smith is partly due to a failure to distinguish between two quite different types of case. Suppose we have a set of values for $x$ and $y$, and that the values of $y$ depart irregularly from a parabolic or other simple function of $x$, and that it is our intention to find the constants of such a function which best represents the relation. Two distinct cases may arise. In one case the given values of $x$ and $y$ may be exact, or at any rate only subject to errors of measurement which are negligible compared to the departures of the individual $y$'s from the corresponding values of $f(x)$. In other words, the data accurately represent the facts, which do not actually fit any simple formula of the proposed type. We may nevertheless, for various reasons, desire to find, say, a parabolic formula which *most nearly* fits the facts. Our problem is to find the constants in an admittedly approximate relation which will depart from the truth by the smallest practicable amount at any value of $x$ taken at random. In such a case any of the given values of $y$ is as important as any other and the criterion adopted by Mr Smith is clearly the only logical one to apply. The other type of case, which is of much greater importance to the experimental physicist, is when the departures of $y$ from the corresponding values of any simple function of $x$ are of the same order as the errors of observation, that is to say, when one or more functions can be found which represent the experimental data to within the accuracy of the observations. This case arises frequently in physical measurements, as for example, in determining the resistance-temperature relation of a resistor, the e.m.f.-temperature relation of a thermocouple, and so on.

The classical principles for the reduction of such observations treat this case as exactly equivalent to the former. It is assumed that every observation is of equal importance and must be weighted in accordance with a prescribed mathematical procedure—in brief, the method of least squares. I recall that when learning physics in my school days, every figure, once it was recorded in the notebook, became sacrosanct: it must not be rubbed out, or amended, or one's reputation for honesty was seriously compromised. To use one's intelligence in discarding obviously anomalous results was regarded as "faking" the answer. This principle is still applied in the teaching of experimental science, and is based on the proposition that as every observation is made in ignorance of any predisposing cause of error (otherwise it wouldn't be made) there is no reason for considering one observation more or less reliable than another: the probability of being right is the same for all. This proposition is false. Probability always means, or should mean, probability in the light of the available evidence. At the moment when each observation is made, the probability of its being correct is the same as for any other of the series, but this is no longer the case when the whole series is completed and available for inspection. The trend of the observations, as a whole, constitutes additional evidence which affects the estimate of probability to be applied to any one observation. It provides logical justification for cutting out observations which are abnormal before proceeding to the formal reduction. As to what constitutes an abnormal observation there is no formal rule which can replace the intelligence of the experimenter: his criterion will be based on experience of the particular kind of measurement which is in question and accumulated knowledge of the behaviour of the apparatus; but a rough general rule, to be used with, not as a substitute for, intelligence is that any single observation which, if admitted, would appreciably alter the result from that obtained by leaving it out should be discarded. If too many of the observations turn out to be worthless when separately tested in this way we do not mend matters by leaving them in: the only remedy in such a case is to do the job over again with increased care or by improved methods. But every experimenter knows that even when all reasonable care has been taken anomalous observations tend to crop up. To reject these is not faking, but is merely the use of intelligence and a proper interpretation of probability, and any series of observations should have such outlying members pruned off before it is subjected to mathematical reduction. The experimenter who fails to do this, in deference to the classical misapplication of statistical theory to experimental observations, is merely permitting the intelligence which he has presumably brought to bear on the experimental part of the work to abdicate when he comes to the arithmetic.

A series of observations, cleaned of obvious errors, will never present the kind of anomalous distribution to which one would fear to apply Mr Smith's mid-course criterion.

I would submit, however, that in the case now being considered, where the irregularities to be smoothed out are of the same order as observational errors, there is no theoretical virtue in any particular method of finding the constants of the function, and that the only relevant consideration is the time involved in the computations. The difference in the function as evaluated by any reasonable methods represents the inherent ambiguity of the given data, and it is unjustifiable to assume that any one of the perfectly good curves which can be made to fit it is *truer* than another. One of the crimes to be laid at the door of the statisticians in their treatment of experimental observations is that they have encouraged the belief that by suitable mathematical manipulation of irregular observations the true value of some quantity, or of the constants in some relation, can be established within limits of uncertainty which are in fact illusory. Ample proof of this is afforded by the literature: how often do we find different experimenters determining the same thing and disagreeing by amounts in excess of the sum of their estimated uncertainties as computed by the classical formulae. My experience is that any series of observations can be smoothed graphically with as much accuracy as the observations justify. Where mere smoothing is not enough and we want the constants of a functional relation, we cannot do better than employ whatever method gives these with the minimum amount of calculation. It seems clear, on the evidence, that the mid-course method is the best in this respect and that the least-squares method is far and away the worst.

Author's reply. I have not attempted in the paper to deal with the observations, but have assumed that there is no reason to believe that any of them are mistakes and that the differences or scattering are such as are usually to be expected in the kind of experiment considered. Though under these conditions it would be very unusual to make as many as fifty measurements, yet circumstances can be imagined where the best value from fifty observations $A + dA$ and one $A - dA$ would be $A$. The point I would make is that the observer ought to consider the most important factors in each problem and not apply a stock method in a routine manner. I am interested to hear that Lord Rayleigh was advised to forget the theory of least squares, and I should like to see the same advice given generally to young physicists today.

In reply to Mr Adams: I should not hesitate to apply the mid-course method to any satisfactory set of quantities whenever I believed the result required was likely to lie in a central position between the extreme observations. I can imagine circumstances in which the method could reasonably be applied to vital statistics and kindred magnitudes.

I am in agreement with much that Mr Guild says, but I have not reduced the question of the rejection of observations to a definite set of rules in the way he has done. When there is no question of mistakes the problem seems to me to contain a number of difficulties. The fact that Mr Guild's work is regularly of very great accuracy may be considered a recommendation of the views he puts forward.

I agree with Mr Guild that many workers will find the small amount of arithmetic involved the most important property of the mid-course method. Its rapidity is of course due to the fact that only a few observations enter into the calculations. High-order curves are not often wanted, but I have recently compared various methods of computing the coefficients and have found the following general method particularly convenient. I assume that the polynomial to be obtained is of order $2n$, so that $2n + 2$ outstanding observations are utilized. Let the polynomial be

$$f(x) = a_1 x^{2n} + a_2 x^{2n-1} + a_3 x^{2n-2} + \dots + a_{2n+1}.$$

I find from the odd observations a linear matrix $Y$ and square matrices $S_q$ and $R_q$, each of order $n$. Their definitions are $Y = (Y_0, Y_1, Y_2, \dots Y_{n-1})$,

$$S_q = \begin{pmatrix} s_q & s_{q+1} & s_{q+2} & \cdots \\ s_{q-1} & s_q & s_{q+1} & \cdots \\ s_{q-2} & s_{q-1} & s_q & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}, \quad R_q = \begin{pmatrix} r_q & r_{q+1} & r_{q+2} & \cdots \\ r_{q-1} & r_q & r_{q+1} & \cdots \\ r_{q-2} & r_{q-1} & r_q & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix},$$

where

$$\begin{aligned}(x - x_1)(x - x_3) \dots (x - x_{2n+1}) &= x^{n+1} + s_1 x^n + s_2 x^{n-1} + \dots + s_{n+1} \\ &= x^{n+1}\{1 + r_1 x^{-1} + r_2 x^{-2} + \dots\}^{-1},\end{aligned}$$

$$Y_u = \sum_{v=0}^{n} \frac{y_{2v+1}\, x^u{}_{2v+1}}{(x_{2v+1} - x_1)(x_{2v+1} - x_3) \dots (x_{2v+1} - x_{2n+1})}.$$

From the even observations similar matrices $Y'$, $S_q'$, $R_q'$ are found. Then the first $n$ of the coefficients $a$ are given by

$$(a_1\, a_2\, a_3 \dots a_n) = (Y' S_0' - Y S_0)(R_0 S_n - R_0' S_n')^{-1},$$

the succeeding $n$ coefficients by

$$\begin{aligned}(a_{n+1}\, a_{n+2} \dots a_{2n}) &= (a_1\, a_2 \dots a_n) R_0 S_n + Y S_0 \\ &= (a_1\, a_2 \dots a_n) R_0' S_n' + Y' S_0',\end{aligned}$$

and the constant term $a_{2n+1}$ and the error $\epsilon$ by

$$y_u + y_v - 2a_{2n+1} = \{a_1 x_u{}^{2n} + a_2 x_u{}^{2n-1} + \dots + a_{2n} x_u\} + \{a_1 x_v{}^{2n} + a_2 x_v{}^{2n-1} + \dots + a_{2n} x_v\},$$

$$y_u - y_v \pm 2\epsilon = \{a_1 x_u{}^{2n} + a_2 x_u{}^{2n-1} + \dots + a_{2n} x_u\} - \{a_1 x_v{}^{2n} + a_2 x_v{}^{2n-1} + \dots + a_{2n} x_v\},$$

where $x_u$, $y_u$ and $x_v$, $y_v$ are any two observations taken one from the even and the other from the odd set. Formulae for the polynomial of order $2n - 1$ may be derived by putting $x_{2n+2} = \infty$.

519.281

# THE DETERMINATION OF A PARABOLIC FORMULA TO REPRESENT A SERIES OF OBSERVATIONS

By J. H. AWBERY, B.A., B.Sc., F.Inst.P., Physics Department,
National Physical Laboratory, Teddington, Middlesex

*Received February* 16, 1934. *Read May* 18, 1934.

*ABSTRACT.* When the constants $a$, $b$ and $c$ in the formula $y = a + bx + cx^2$ are to be determined from simultaneous values of $x$ and $y$, the best-known methods involve the addition of a number of equations, which are then solved to give the desired constants. A large number of significant figures must be retained in the calculation. It is pointed out that if one constant can be determined separately from the other two, the arithmetic can be greatly simplified. It is frequently of advantage to determine first the constant $c$, since this is the constant needed to correct mean values to instantaneous values. It can be easily found by the method of divided differences which, in contrast to the more familiar methods, reduces the number of significant figures to the minimum permissible. If the observations are very irregular, the method in its crudest form gives a very poor estimate of $c$, but a modification is described which adds little to the labour, whilst improving the accuracy very considerably.

Numerical examples are given, and the results by the new method compared with those obtained by the methods of least squares and zero sum.

## § 1. INTRODUCTION

$c, x, y$

The parabolic equation $y = a + bx + cx^2$ is of common occurrence in physics and engineering, either as a rigorous equation, or as one to represent a series of results empirically. Its utility for the latter purpose has been stressed recently by various authors. For example, Tuttle and Satterly claim that in many cases it is preferable as an interpolation formula to the commonly adopted form $y = ax^n$.

When such a formula is to be adopted, a variety of methods are in use for determining the constants $a$, $b$ and $c$ from the observations. The authors just cited, for example, suggest drawing a curve to represent the observations, and, selecting three points from the curve, to provide three equations for $a$, $b$ and $c$. Clearly this method is suitable only for relatively low accuracy. When greater precision is necessary, the two most familiar devices for obtaining three equations are the method of least squares and that of zero sum. The former involves a great deal of labour before the normal equations can be formed, whilst both methods demand the use of a considerable number of significant figures in the calculations, in order to obtain even a moderate number in the values of $a$, $b$ and $c$.

To illustrate this, we may work a very simple example by the method of zero sum. It is desired to calculate $a$, $b$ and $c$ in the equation $y = a + bx + cx^2$ when the following values are given:

| | | | | | |
|---|---|---|---|---|---|
| $x = 1$ | 2 | 3·5 | 4 | 5 | 8 |
| $y = 1$ | 2 | 4 | 7 | 14 | 22 |

The first step is to write out the six quadratic equations, and to reduce them to three by adding in pairs. The result is

$$\left.\begin{aligned} 3 &= 2a + 3b + 5c \\ 11 &= 2a + 7{\cdot}5b + 28{\cdot}25c \\ 36 &= 2a + 13b + 89c \end{aligned}\right\} \qquad \ldots\ldots(1).$$

Subtracting each of these from the following one eliminates $a$ and yields

$$\left.\begin{aligned} 8 &= 4{\cdot}5b + 23{\cdot}25c \\ 25 &= 5{\cdot}5b + 60{\cdot}75c \end{aligned}\right\} \qquad \ldots\ldots(2).$$

If, now, each of these is divided by the coefficient of ($b$), with only slide-rule accuracy, they become respectively

$$\left.\begin{aligned} 1{\cdot}775 &= b + 5{\cdot}17c \\ 4{\cdot}545 &= b + 11{\cdot}03c \end{aligned}\right\} \qquad \ldots\ldots(3),$$

whence by subtraction $$2{\cdot}770 = 5{\cdot}86c \qquad \ldots\ldots(4),$$

or $$c = 0{\cdot}473.$$

With this value in (3), $b = -0{\cdot}67$; and then from (1) $2a = 2{\cdot}64$ or $a = 1{\cdot}32$.

The above approximate solution was carried out with a slide-rule, and it is now possible to estimate the accuracy really required. Since $y$ is given only to the nearest integer, it is desirable that the formula should give the first decimal place. Now, when $y$ is being calculated for $x = 8$, the contributions of the three terms are approximately $1{\cdot}3$, $-5{\cdot}3$ and $-30{\cdot}3$, and $a$ should therefore be known to 2, $b$ to 2 and $c$ to 3 significant figures. Thus as regards $c$, equation (4) has only been taken to exactly sufficient places; since the coefficient 5·86 is obtained by subtraction, it may be in error by 1 unit in the last place, and therefore one more figure would be desirable. Again, in finding $b$ from equation (3), we subtract $11{\cdot}03c$ from 4·545, i.e. we subtract 5·22 from 11·03. Clearly, to be sure of the second decimal place, i.e. to know $b$ to two figures, it would be preferable to know the value of the quantity $11{\cdot}03c$ to one more figure. Similarly, the second significant figure of $a$ would be more sure if $c$ were known to one more figure*.

It is apparent that when any constant is required with greater precision than these, as is usually the case, the calculations are outside the range either of a slide-rule or of four-figure logarithms.

In the method of least squares, the amount of arithmetic required is, of course, greater even than in the method of zero sum.

## § 2. PROPOSED METHOD

It is clear from the foregoing that a simpler method of determining the constants in the equation $y = a + bx + cx^2$ would be of value. Much of the labour in the standard methods arises from the fact that the equations are added, numbers with many significant figures being thus introduced, and all the constants are determined simultaneously.

---

* As a matter of fact a more exact calculation shows that the values $a, b, c = 1{\cdot}32, -0{\cdot}67$ and 0·473 should be 1·30, $-0{\cdot}65$ and 0·471, so that slide-rule accuracy introduces perceptible errors even in this case.

If any one of the constants can be determined before the others, the problem becomes much simpler. Thus if $c$ is known first, it is easy to plot $(y - cx^2)$ against $x$ and so to determine $a$ and $b$ as the intercept and slope of a straight line; if $b$ is known, $(y - bx)$ is a linear function of $x^2$, whilst if $a$ is determined first, then $(y - a)/x$ is linear against $x$. (The reference to plotting does not imply that graphical methods are necessary. There are plenty of easy arithmetical methods for determining the two constants in a linear formula.)

For several reasons, the constant $c$ is the most useful to determine first. It is often the coefficient of a small term, and the arithmetic involved in preparing the linear function to determine $a$ and $b$ is easily carried out. Moreover, as will be shown elsewhere, if mean values have been measured for a quantity which varies with $x$, it is in any case desirable to know $c$ in order to correct these values to instantaneous values.

Fortunately $c$ is not only the most useful constant to determine first; it is also the easiest. For consider three values of $x$ and the corresponding values of $y$, say

$$y_1 = a + bx_1 + cx_1^2,$$

$$y_2 = a + bx_2 + cx_2^2,$$

$$y_3 = a + bx_3 + cx_3^2.$$

By taking the differences of the $y$-values and those of the $x$-values, and dividing, the first divided differences of $y$ are found. They are seen to be given by

$$(y_2 - y_1)/(x_2 - x_1) = b + c\,(x_1 + x_2) \quad \text{and} \quad (y_3 - y_2)/(x_3 - x_2) = b + c\,(x_2 + x_3).$$

The second divided difference is obtained by subtracting one first difference from the other and dividing by $(x_3 - x_1)$, the range of values of $x$ concerned in its formation. From the above equations it is immediately seen that the second divided difference is simply equal to $c$.

A numerical example will make the matter clear.

Table 1. Divided differences of a function

| $x$ | $y$ | Differences of | | First divided difference of $y$ | Second divided difference of $y$ | | |
|---|---|---|---|---|---|---|---|
| | | $x$ | $y$ | | Numerator | Denominator | Value |
| 1 | 9 | | | | | | |
| | | 1 | 6 | 6 | | | |
| 2 | 15 | | | | 9 − 6 or 3 | 4 − 1 or 3 | 1 |
| | | 2 | 18 | 9 | | | |
| 4 | 33 | | | | 12 − 9 or 3 | 5 − 2 or 3 | 1 |
| | | 1 | 12 | 12 | | | |
| 5 | 45 | | | | 15 − 12 or 3 | 7 − 4 or 3 | 1 |
| | | 2 | 30 | 15 | | | |
| 7 | 75 | | | | 22 − 15 or 7 | 12 − 5 or 7 | 1 |
| | | 5 | 110 | 22 | | | |
| 12 | 185 | | | | 28 − 22 or 6 | 13 − 7 or 6 | 1 |
| | | 1 | 28 | 28 | | | |
| 13 | 213 | | | | | | |

### § 3. A NUMERICAL EXAMPLE

To show the operation of the method, the following table of corrections to the constant-volume helium thermometer will be taken, and a parabolic formula deduced to represent the results. The first column gives the temperature, the second the correction in thousandths of a degree. The working is shown in the later columns, the last of which should be a linear function of $x$. A plot of the figures verifies this, and consequently shows that the parabolic form is suitable for representing the data.

Table 2

| $x$ | $y$ | $\Delta y$ | $\Delta^2 y$ | $cx^2$ | $y - cx^2$ |
|---|---|---|---|---|---|
| − 183 | 19 | | | 12·8 | 6·2 |
| | | − 0·151 | | | |
| − 150 | 14 | | + 0·00037 | 8·6 | 5·4 |
| | | − 0·120 | | | |
| − 100 | 8 | | + 0·00020 | 3·8 | 4·2 |
| | | − 0·100 | | | |
| − 50 | 3 | | + 0·00040 | 1·0 | 2·0 |
| | | − 0·060 | | | |
| 0 | 0 | | + 0·00040 | 0·0 | 0·0 |
| | | − 0·020 | | | |
| 50 | − 1 | | + 0·00040 | 1·0 | − 2·0 |
| | | + 0·020 | | | |
| 100 | 0 | | + 0·00040 | 3·8 | − 3·8 |
| | | + 0·060 | | | |
| 150 | 3 | | + 0·00040 | 8·6 | − 5·6 |
| | | + 0·100 | | | |
| 200 | 8 | | + 0·00040 | 15·2 | − 7·2 |
| | | + 0·140 | | | |
| 250 | 15 | | + 0·00020 | 23·8 | − 8·8 |
| | | + 0·160 | | | |
| 300 | 23 | | + 0·00060 | 34·3 | − 11·3 |
| | | + 0·220 | | | |
| 350 | 34 | | + 0·00020 | 46·7 | − 12·7 |
| | | + 0·240 | | | |
| 400 | 46 | | + 0·00060 | 61·0 | − 15·0 |
| | | + 0·300 | | | |
| 450 | 61 | | | 77·2 | − 16·2 |
| | Mean $c = 0{\cdot}000381$ | | | | |

The time occupied in forming this table, including the copying out of the first two columns, was 9 minutes, of which less than half sufficed to determine $c$.

To determine $a$ and $b$, a variety of methods is available, and we will select that of zero sum. Adding the upper halves of the first and last columns gives $12{\cdot}0 = 7a - 333b$, and adding the lower halves gives $-76{\cdot}8 = 7a + 2100b$. By difference, $2433b = -88{\cdot}8$ or $b = -0{\cdot}0365$, and then $7a = -0{\cdot}17$ or $a = -0{\cdot}02$.

The time occupied in this determination of $a$ and $b$ was $5\frac{1}{2}$ minutes, making a total of $14\frac{1}{2}$ minutes for the whole process of determining the parabolic equation

$$y = 0{\cdot}02 - 0{\cdot}0365x + 0{\cdot}000381x^2.$$

To solve the same problem by the method of zero sum, it is first necessary to write down the fourteen equations of which the first is

$$19 = a - 183b + 33489c,$$

and to divide these into three equal groups of $4\frac{2}{3}$ equations each*. These groups are then added, to form the three normal equations

$$\begin{aligned} 44 &= a - 483b + 68{,}490c, \\ 15 &= a + 583{\cdot}3b + 95{,}833c, \\ 174 &= a + 1666{\cdot}7b + 616{,}667c, \end{aligned}$$

a process which was found to occupy 8 minutes.

If these are solved with merely slide-rule accuracy (which is not really sufficient), they give $a = -0{\cdot}01$, $b = -0{\cdot}0370$ and $c = 0{\cdot}000382$, the time occupied being 11 minutes, making a total of 19 minutes. The difference in the results by the two methods is insignificant, although the time-difference is over 30 per cent.

In the solution by least squares, even the formation of the normal equations occupied 31 minutes, and their solution 16 minutes.

When the results calculated by the various methods are compared with the original data, each method shows a maximum deviation of 0·6 of a unit, though these deviations all occur at different values of the temperature $x$.

## § 4. REFINEMENT OF THE METHOD

In the example just taken, the differences of both the first and second orders were found by subtracting successive entries in the table. This was justifiable because the twelve estimates of $c$ did not vary greatly, but in general this would not be so. This procedure does not weight the observations equally, since every $y$-value except the first and last enters into the calculation of two first differences. Consequently the first two and last two entries are weighted differently from each other and from the rest of the entries. Again, if the observations are scattered at all widely about the mean parabola representing them, some of the first and second differences will be reversed in sign with respect to the rest. For example, if the trend of the observations is such as to demand a parabola with its axis vertical and vertex downwards, then a single high observation lying between two normal or low ones introduces a value of $c$ representing a parabola with vertex upwards. Again, if three observations are very close together in their $x$-values, they may give quite large divided first differences, and these will exercise undue weight in the second differences.

For these reasons, it is desirable to form the first differences by dividing the table into two equal halves, ignoring the middle observation if there are an odd number, and subtracting the values in corresponding positions in the two halves. In this way every observation which is used at all is used just once. The same principle is then to be extended to the second differences. Thus, supposing there are $2n$ observations the first divided differences will be

$$\frac{y_{n+1} - y_1}{x_{n+1} - x_1}, \quad \frac{y_{n+2} - y_2}{x_{n+2} - x_2}, \ \ldots,$$

---

* A fraction such as two-thirds of an equation is obtained by multiplying every coefficient by $\frac{2}{3}$. E.g. $\frac{2}{3}$ of the equation $3 = a + 6b + 36c$ is $2 = \frac{2}{3}a + 4b + 24c$.

i.e. $$b + c\,(x_{n+1} + x_1), \quad b + c\,(x_{n+2} + x_2), \ \dots$$

and the second will proceed analogously.

To illustrate the improvement effected by this refinement of the procedure, we may take an example of scattered data from Merriman's *Least Squares*, p. 140. The observations are shown and are treated by the cruder method in table 3.

Table 3

| $x$ | $y$ | Divided first differences | Divided second differences |
|---|---|---|---|
| 28 | 11·71 | | |
| | | 1·19/38 or 0·0313 | |
| 66 | 12·90 | | + 0·0014/145 or + 9·7 × $10^{-6}$ |
| | | 3·50/107 or 0·0327 | |
| 173 | 16·40 | | + 0·0153/182 or + 84·1 × $10^{-6}$ |
| | | 3·60/75 or 0·0480 | |
| 248 | 20·00 | | − 0·0040/125 or − 32·1 × $10^{-6}$ |
| | | 2·20/50 or 0·0440 | |
| 298 | 22·20 | | − 0·0288/152 or − 189·5 × $10^{-6}$ |
| | | 1·55/102 or 0·0152 | |
| 400 | 23·75 | | + 0·0105/207 or + 50·7 × $10^{-6}$ |
| | | 2·70/105 or 0·0257 | |
| 505 | 26·45 | | + 0·0034/148 or + 23·0 × $10^{-6}$ |
| | | 1·25/43 or 0·0291 | |
| 548 | 27·70 | | Mean − 9·0 × $10^{-6}$ |

Each entry in the last column is an estimate of the constant $c$, and it is obvious that the mean is quite unreliable. The most favourable complexion is given to the figures by taking first the mean of the three estimates − 189·5, + 84·1 and + 50·7 (i.e. − 18·2) and secondly the mean of the other three (viz. + 0·2) and comparing these means, but it certainly gives no confidence in the final result.

If now the half-table method is adopted, the calculation takes the form given in the next table.

Table 4

| $x$ | $y$ | Divided first differences | Divided second differences |
|---|---|---|---|
| 28 | 11·71 | | |
| 66 | 12·90 | | |
| 173 | 16·40 | | |
| 248 | 20·00 | | |
| 298 | 22·20 | 10·49/270 or 0·03886 | |
| 400 | 23·75 | 10·85/334 or 0·03249 | |
| 505 | 26·45 | 10·05/332 or 0·03027 | − 0·00859/(207 + 145) or − 0·0000244 |
| 548 | 27·70 | 7·80/300 or 0·02566 | − 0·00683/(148 + 182) or − 0·0000207 |

The two estimates of $c$ in the last column are much closer together than the two estimates obtained previously by grouping three observations together to obtain a mean, and in fact it seems probable that two-figure accuracy can be justifiably claimed in the value of $c$ now deduced, although three would be used in practice, to avoid errors due to rounding off.

A second important point is that the value − 22·6 × $10^{-6}$ deduced for $c$ in table 4 not only is apparently more reliable, but is nearly three times as large as that deduced in table 3. This larger value is much nearer to the value $-23{\cdot}1 \times 10^{-6}$* given by the method of least squares and to the value $-32{\cdot}2 \times 10^{-6}$ given by the method of zero sum; in view of the differences between the values due to these older methods, it is to be presumed that the observations depart so far from a parabola that any value between $-19$ and $-32$ is as "true" as any other.

## § 5. DETERMINATION OF THE CONSTANTS *a* AND *b*

Attention has so far been mainly devoted to the determination of the constant $c$, which is all that would be required in correcting mean values to instantaneous ones. If the whole equation is to be found, the constants $a$ and $b$ will be required, and many methods are available for determining them.

In one of the examples given above the quantity $y - cx^2$ was formed and zero sum then utilized. This is a convenient method since it easily gives two equations for the two unknowns, one of which is immediately eliminated by subtraction. An equally simple method, and one which gives a visible check on whether a parabolic equation is in fact satisfactory, is to form as before the quantity $y'$
$y'$ which $= y - cx^2 = a - bx$. The half-table method is then used to determine $b$, which is given as the ratio between the differences of $y'$ and of $x$. When $b$ has been determined, an estimate of $a$ is obtained from each observation by subtracting $bx$ from the previously tabulated quantity $y'$. The deviations of these estimates of $a$ from their mean are identically equal to the deviations of the calculated values of $x$ from the observed values, and this is a valuable feature of this method of determining $a$ and $b$.

Yet another method of determining $a$ and $b$ is available. In the process of finding $c$, first divided differences are formed and are equal to $b - c\,(x_r + x_s)$. Hence by subtracting the appropriate value of $c\,(x_r + x_s)$ from each divided difference, a number of estimates of $b$ are obtained.

These methods are illustrated for the problem from Merriman just dealt with, table 5 showing the half-table method.

Table 5

| $x$ | $y$ | $x^2$ | $cx^2$ ($c = 22{\cdot}6 \times 10^{-6}$) | $y'$ | $b$ | $bx$ ($b = 0{\cdot}0434$) | $a$ or $y' - bx$ |
|---|---|---|---|---|---|---|---|
| 28 | 11·71 | 780 | 0·02 | 11·73 | | 1·22 | 10·51 |
| 66 | 12·90 | 4360 | 0·10 | 13·00 | | 2·86 | 10·14 |
| 173 | 16·40 | 2930 | 0·68 | 17·08 | | 7·51 | 9·57 |
| 248 | 20·00 | 61500 | 1·39 | 21·39 | | 10·76 | 10·63 |
| 298 | 22·20 | 88800 | 2·01 | 24·21 | 12·48/270 or 0·0412 | 12·93 | 11·28 |
| 400 | 23·75 | 16000 | 3·61 | 27·36 | 14·36/334 or 0·0430 | 17·36 | 10·00 |
| 505 | 26·45 | 255020 | 5·76 | 32·21 | 15·13/332 or 0·0456 | 21·92 | 10·29 |
| 548 | 27·70 | 300300 | 6·78 | 34·48 | 13·09/300 or 0·0436 | 23·78 | 10·70 |
| | | | | | Mean 0·0434 | Mean | 10·39 |

---

* The result quoted in Merriman's book is not the least-squares result to the problem worked out here. It is the conditioned solution forced to pass through the point representing the surface temperature.

The first-divided-difference method is set out in table 6.

Table 6

| $x$ | First divided differences | $(x_r + x_s)$ | $c\,(x_r + x_s)$ | $b$ | $bx$ ($b = 0{\cdot}0446$) | $y - cx^2$ | $a$ |
|---|---|---|---|---|---|---|---|
| 28 | | | | | 1·25 | 11·73 | 10·48 |
| 66 | | | | | 2·94 | 13·00 | 10·06 |
| 173 | | | | | 7·72 | 17·08 | 9·36 |
| 248 | | | | | 11·05 | 21·39 | 10·34 |
| 298 | 0·03886 | 326 | − 0·0074 | 0·0463 | 13·29 | 24·21 | 10·92 |
| 400 | 0·03249 | 466 | − 0·0105 | 0·0430 | 17·84 | 27·36 | 9·52 |
| 505 | 0·03027 | 678 | − 0·0153 | 0·0456 | 22·51 | 32·21 | 9·70 |
| 548 | 0·02566 | 796 | − 0·0180 | 0·0437 | 24·42 | 34·48 | 10·06 |
| | | | Mean | 0·0446 | | Mean | 10·06 |

We have thus a number of alternative equations, each purporting to represent the data:

From least squares $10{\cdot}21 + 0{\cdot}0442x - 23{\cdot}1 \times 10^{-6}x^2$;

from zero sum $9{\cdot}75 + 0{\cdot}0495x - 32{\cdot}2 \times 10^{-6}x^2$;

from the difference method, using the half table to determine $b$,

$10{\cdot}39 + 0{\cdot}0434x - 22{\cdot}6 \times 10^{-6}x^2$;

from the difference method, using the first divided differences to determine $b$,

$10{\cdot}06 + 0{\cdot}0446x - 22{\cdot}6 \times 10^{-6}x^2$.

The accuracy of these various equations can be estimated by calculating $y$ from each at the given values of $x$, and noting the deviations between the observed and calculated values of $y$. The main results of such a comparison are shown in table 7, from which it will be seen that in point of fidelity to the data there is little to choose between the various methods.

Table 7

| | Least squares | Zero sum | Difference method | |
|---|---|---|---|---|
| | | | Half table | First differences |
| Average deviation | 0·40 | 0·43 | 0·39 | 0·40 |
| Maximum deviation | 0·87 | 0·96 | 0·82 | 0·86 |
| Sum of squares of deviations | 1·84 | 2·27 | 1·85 | 1·90 |

## DISCUSSION

Dr Ferguson. It is all very well to discourse opulently of reductions made in five or ten minutes with the aid of mechanical computers, but those of us whose normal equipment consists of a thirty-centimetre boxwood rule, a few pieces of squared paper of doubtful rectangularity, and a well-thumbed copy of Bottomley, do not disdain more primitive methods which will achieve a result in something less than half an hour.

In handling an equation of the type considered by Mr Awbery,

$$y = a + bx + cx^2,$$

if we plot the best $xy$ curve, take a point $x_0 y_0$ thereon, and transfer the origin to this point we can, by plotting the new values of $y/x$ as ordinate against $x$ as abscissa, obtain a straight line the slope and intercept of which determine $c$ and $b$.

Having found these constants, from each pair of original values of $y$ and $x$ we can determine a value of $a$ by using the original equation.

This method of reducing a four-term equation to what is practically a three-term form by transferring the origin to a point on the curve is certainly restrictive, but it is simple, of fairly general application, and it obeys the pragmatic test.

Mr T. SMITH. From an examination of the outstanding errors in a number of problems when formulae have been fitted by least squares, by the method of this paper, and by others, the conclusion emerges that this method tends to give solutions which approach very nearly—certainly more closely than other methods included in the comparison—to the least-squares solutions but with much less labour than the latter involve. The method ought therefore to prove particularly valuable where a simple, rapid and reliable method in which weight is given to all the observations is required.

AUTHOR'S reply. I sympathize with Dr Ferguson's plea for the man whose study is not well equipped with aids to computing, but surely the fact that the squared paper is of doubtful accuracy should be an argument against, rather than for, graphical methods? I am an advocate of reducing numerical work where possible to the determination of the parameters of a straight line, but I prefer to determine them arithmetically, using the eye only as a safeguard against the adoption of an unsuitable formula, or the introduction of a gross mistake.

My objections to Dr Ferguson's procedure in the particular case of a quadratic formula are centred around the selection of a point $(x_0, y_0)$ through which the curve is forced to pass. I feel that this selection may bias the whole curve, and further that the selection may be no simple matter. Consider the example in table 2 of Mr Smith's paper*. Here, values of $v$ range from 2·97 to 3·20 and are given to four decimal places. If they are plotted so that 1 cm. represents 0·01 of a unit (a curve 9 inches wide), then to estimate the fourth figure in $v_0$ requires reading the curve to the nearest tenth of a millimetre. I presume Dr Ferguson would surmount this difficulty by plotting, not the observations, but their deviations from a straight line passing through the third and eighth points; but if so, there is quite considerable labour to be done before he starts the main process.

---

* Page 566 of this volume.

# REVIEWS OF BOOKS

## *The General Principles of the Quantum Theory*, by G. Temple. Pp. viii + 120. (London: Methuen & Co., Ltd.) 3*s*.

Methuen's monographs on physical subjects, to which series this booklet belongs, are mostly excellent value, and the present one is certainly no exception to the rule. It introduces the student to all the main conceptions of modern quantum theory, and does not omit to show him how to work the mathematical apparatus which has been adopted in this branch of theoretical physics.

To do all this in the space available would be impossible without tremendous compression, and the book is therefore difficult reading. Of necessity, the arguments are mainly expressed in mathematical symbolism, and explanations as to how one equation follows from another have been kept to a minimum.

Nevertheless, it can be read by any conscientious student, and he can assure himself that his knowledge is real and not superficial by working the examples. These add greatly to the value of such a work as this, and make it more valuable as an introduction to the subject than works of many times the size with no examples. There are one or two misprints in them, as on p. 11, ex. 2, where the replacement of "are" by "and" makes it look as if the enunciation is incomplete. The other misprints (two on p. 41 and several in chapter 5) are even more trivial.

It is refreshing to find an author who considers that English is good enough for him. For once, we find a book on quantum theory in which "characteristic" is not rendered as "eigen," nor "trace" as "spur."

J. H. A.

## *Wave Mechanics. Advanced General Theory*, by J. Frenkel. Pp. viii + 525. (Oxford: Clarendon Press.) 35*s*.

This is the second volume of a trilogy, of which the first dealt with elementary general theory and the third will treat special problems in detail.

The distinction between the elementary theory of the first volume and the advanced theory of the one under review is very marked. Here the relationship to classical mechanics is developed from the most refined formulation of the latter subject, and the equations of the new mechanics are built up in the first place as generalizations of classical mechanics with operators taking the place of quantities. In this synthesis, the analogy between classical mechanics and geometrical optics on the one hand, and wave mechanics and undulatory optics on the other, is not used at all.

The development by which the operators are represented as matrices is next taken up, this part of the subject being completed on p. 126.

The next two chapters are devoted to the transformation theory. Here we are given two operators $H$ and $K$, representing the energy of a particle in two fields of force with different potential energies, and are required to find what relations exist between the characteristic functions $\psi_H$ and $\psi_K$, and also between the two matrices which represent a given operator $E$, in its relation first to $H$ and then to $K$. The perturbation theory is a special case, where the energies $H$ and $K$ are related so that one is an approximation to the other.

This brings us practically half way through the book, and the remainder is much more difficult reading. It opens with a chapter on "relativistic remodelling and magnetic generalization of the wave mechanics of a single electron." This chapter, of course, introduces the spin effect and is well written to show the way in which the magnetic generalization is indissolubly linked to the relativistic outlook. After it comes the problem of many particles, and the treatment by what is variously known as second quantization, super-quantization or intensity quantization.

This second half of the book contains much that is very recent, and naturally has to deal with problems where the best solution that can be given is known to be incomplete or incorrect. In the main it follows the treatment given by Dirac in his various papers, though there is no slavish adherence to any one line of discussion. It will be seen that the term "wave mechanics" is used in a very wide sense, and includes modern quantum electrodynamics.

As a whole the book is difficult reading. In many ways it is reminiscent of Thomson and Tait's *Natural Philosophy*. Both books adopt an extremely general outlook, though the older one is kinder in giving examples (even trivial ones are helpful) of the sort of system to which any given theorem applies. This is balanced by the greater discursiveness in Frenkel's book, and in particular by the lengthy verbal translations which frequently precede the mathematical investigations.

The books have another similarity—their weight and general unsuitability for armchair reading. In the present case this is due partly to the very thick paper which the publishers have provided. They and the printers are to be congratulated on the excellent production of a book which must have been extremely difficult to set up. The mathematical portions abound with cases where a symbol carries a subscript which itself bears a subscript or a prime or both, whilst double superscripts like $^{1*}$ or $^{\prime\prime *}$ or $^{\prime\prime\ 1}$ are frequent. Despite this complexity only three or four examples of misprints have been noticed, none of them important enough to require citation.

J. H. A.

### *L'Électron Magnétique* (*Théorie de Dirac*), by LOUIS DE BROGLIE. Pp. viii + 316. (Paris: Hermann et Cie, 1934.) 100 fr.

Although Dirac's theory of the electron is usually regarded as one of the high peaks of achievement in mathematical physics, probably few physicists would claim to have more than a vague appreciation of its general character. A full understanding of the theory requires not only a thorough grasp of the spirit of quantum mechanics and of relativity, but also familiarity with a difficult mathematical technique.

Dirac's equations for the electron are arrived at by modifying the usual relativistic wave equation in such a way as to meet a postulated formal requirement that the equations shall be of the first order in time. From the aesthetically satisfying equations so built up may be deduced consequences, in agreement with experiment, as to spin and magnetic moment; so that these appear not as fundamental irrationals, but as particular aspects of a wider scheme.

The essence of the theory is given, with amazing economy of symbolism, in the last chapter of Dirac's *Quantum Mechanics*. De Broglie gives a fuller presentation, leading up by gradual steps from what will be familiar to many physicists to the more recondite theory.

In the first part of the book there is a good survey of the relevant parts of classical quantum theory and of wave mechanics. The second part deals with the Dirac theory, with due reference to the earlier work of Pauli and Darwin. The arguments of Dirac are presented in a manner which will be appreciated by those who wish to gain a clear impression of the nature of the theory without working through the detailed mathematical development which follows. This part of the book concludes with a formidable table of the sixteen hermitic operators involved in the theory, together with expressions for the operators corresponding to mass, charge, current, magnetic and electric moment, and spin.

The third part of the book deals with applications, and the theory of fine structure and of the Landé $g$ factor is fully worked out for the one-electron case. The theory suggests not only magnetic moment and spin, but also an imaginary electric moment, and the possibility of states of negative energy (mass) for the electron. These are briefly discussed, as also questions connected with the lack of symmetry in the appearance of space and time coordinates in the equations.

De Broglie's book is an admirable piece of detailed and lucid exposition which will be of value both to those hesitantly approaching the Dirac theory, and to those requiring a systematic account of it. The book is well printed, and it is adorned with a plate of X-ray spectra and attractive, if irrelevant, head and tail pieces to the chapters showing fragments of line and band spectra.

E. C. S.

*Atomic Theory and the Description of Nature*, by NIELS BOHR. Pp. 119. (London: The Cambridge University Press.) 6*s*. net.

It was a happy thought to bring together five classical essays by Prof. Bohr. Two of them, *Atomic Theory and Mechanics* and *The Quantum Postulate and the Recent Development of Atomic Theory* are familiar to English readers, having appeared in *Nature* in 1925 and in 1927 respectively. The remaining essays are for the first time translated into English. *The Quantum of Action and the Description of Nature* appeared in German in 1929, while *The Atomic Theory and the Fundamental Principles underlying the Description of Nature* was published in Danish in the same year. The fifth essay, entitled the *Introductory Survey*, was published in Danish in 1929, and carries an *Addendum* which was included in the German edition (1931) of a volume which comprised the four articles just named. Description of and comment on these articles are unnecessary, and the convenience of having, in the compass of a small volume, such masterly surveys of the physical thought of the time needs no emphasis.

The historical value of the present volume will be enhanced by the later appearance of a companion volume "containing a number of later essays on the same subject, in which the general point of view is further developed."

A. F.

*The Atom*, by JOHN TUTIN, D.Sc. Pp. 103. (London: Longman's Green & Co., Ltd.) 6*s*. net.

Dr Tutin desires to flype the atom—to give us a structure with the massive protons circulating round a light nucleus. This may seem to some a difficult matter to envisage, but, as Dr Tutin sagely says, "a flywheel with a heavy rim and a light hub is a far more commonplace conception than a flywheel with a light rim and a heavy hub."

He accomplishes his ingenious little essay without any display of mathematical learning, and consequently has not a great deal to say concerning modern spectroscopy. Read critically, the volume will provide a useful exercise in dialectic for a physics student in the honours grade.

A. F.

*Molecular Hydrogen and its Spectrum*, by O. W. RICHARDSON. Pp. xiv + 343. (New Haven: Yale University Press, 1934.)

As the subject of the Silliman Lectures delivered by Prof. Richardson at Yale University in April 1932 and now appearing in book form none, surely, could be more fitting than this complicated spectrum, in the analysis of which the lecturer and his co-workers have taken such an outstanding part extending over about a decade. Having before us a fair-sized book entirely devoted to one spectrum and its emitter, we may well understand the amazement and the doubt expressed by some spectroscopists of sixty to seventy years ago that this complex mass of lines of the so-called "secondary" or "many-lined" spectrum could be associated with the lightest of the chemical elements. The fundamental importance of the spectrum in present-day physics is rooted in the fact that its emitter, $H_2$, the simplest of all molecules, is a structure which the earlier quantum theory entirely failed to interpret but which has been successfully dealt with by the more difficult methods of the new quantum mechanics. Of the many tests which the latter has passed in practical band spectroscopy and other fields none perhaps could be severer than that supplied by a comparison of the theoretical and the observational data for the $H_2$ spectrum.

The authoritative, critical and logically connected account given here will be especially valuable to all students of this difficult subject because the original papers in which the analysis and interpretation of the spectrum have been unfolded are so numerous and so scattered, and also because the book contains a considerable amount of matter which has not hitherto been published and is not intended for publication elsewhere.

The first two of the ten chapters deal with the new quantum-mechanical theory of a pair of nuclei with one electron and with two electrons, in particular the $H_2^+$ ion and the neutral $H_2$ molecule. The general theory of molecular states and spectra is outlined in chapter 3, which deals with transition probabilities, symmetry properties, effects of spins of electrons and nuclei, selection rules, structure of a band system and of a band, and an outline forecast of the $H_2$ spectrum. The next five chapters are devoted to the detailed analysis of the $H_2$ spectrum, one to the singlet systems in the far ultra-violet, two to the singlet systems in the less refrangible regions, and two to the triplet systems. The subjects treated in chapter 9 include molecular energy functions, intensity-distribution, uncoupling phenomena, specific heat of hydrogen, the Raman spectrum of $H_2$, and the $H_2$ continua. The final chapter comprises a section on the theory of the excited states of $H_2$, a discussion of the experimentally established electronic levels, an account of the spectral variations with different conditions of excitation, and a most valuable and comprehensive table of numerical constants for all the known electronic states of $H_2$ and $H_2^+$.

It is particularly fortunate that this masterly work has appeared just when the results of the analysis of the $H_2$ spectrum are being used in connexion with the analyses of the spectra of the isotopic molecules HD and $D_2$ (i.e., $H^1H^2$ and $H_2^2$) which, of course, were not recorded at the time Prof. Richardson delivered these Silliman lectures and wrote them for publication in book form.

W. J.

*Exploring the Upper Atmosphere*, by Dorothy Fisk. Pp. 166. (London: Faber & Faber, Ltd.) 6*s.* net.

One of the most fascinating stories of the century is that of the growth of our knowledge concerning the upper atmosphere; it presents an almost incredible combination of Jules Verne-like fantasies, of human daring, modern atomic physics and recondite technical equipment.

The scope of Miss Fisk's essay may be adequately indicated by a summary of her chapter-headings—Ballooning, Through the troposphere to the lower stratosphere, Sound as an explorer, Ultra-violet rays and the ozone layer, The paths of radio-waves, Projectiles in space, Polar lights and cosmic rays.

The book is intended to appeal to the layman. The author writes in a clear and light-hearted style and has produced a useful elementary introduction to a difficult subject.

A. F.

*Ocean Waves*, by Vaughan Cornish. Pp. xv + 163. (London: The Cambridge University Press.) 10*s.* net.

Physics is not to be regarded as primarily an observational science, and its devotees not infrequently show themselves to be singularly lacking in those powers of rapid, complete and accurate observation which form an essential part of the equipment of a successful naturalist. "I didn't notice it" comes as naturally from the lips of a physicist as from the most Hegelian of metaphysicians.

Dr Vaughan Cornish has provided us with a very valuable study of the size and speed of ocean waves in relation to wind velocity, of waves in sand and snow, and of tidal bores and allied phenomena; and he has demonstrated very clearly what can be accomplished by keen and careful observation plus a certain amount of measurement made with the simplest tools. He has adorned his book with a series of most beautiful photographs of the waves which he describes so lucidly and enthusiastically.

Dr Harold Jeffreys can always be depended on to be clear and terse; the appendix which he contributes is the best short elementary survey known to the present writer of certain aspects of surface wave phenomena.

A. F.

### *L'Excitation Électrique des Tissus*, by A. M. MONNIER. Pp. xvi + 326. (Paris: Hermann et Cie, 1934.) 85 fr.

To explain the wide range of known excitability phenomena the author makes the following two fundamental assumptions: (*a*) a tissue will attain a threshold excitation when the excitatory process $E$ reaches a fixed value $E_0$; (*b*) $E$ is related to the chronaxie $\tau$, the stimulating current $I$ and the time $t$ by the equation

$$a\tau\frac{d^2E}{dt^2}+b\frac{dE}{dt}+\frac{c}{\tau}E=\frac{dI}{dt},$$

where $a$, $b$, $c$ are constants independent of the nature of stimulus or tissue.

Experimental results are compared in detail with the solution of this equation for over a score of different functions of $I$ with respect to $t$. In nearly all cases there is good qualitative agreement, but it is doubtful whether the lack of quantitative correspondence justifies so minute a treatment. A significant fraction of the 300 pages is virtually a restatement, since the quantitative details are too remote from fact to be of importance. On the other hand this very detail, supported by a full bibliography, constitutes a valuable source of reference to the temporal aspects of tissue-excitation.

The author approaches the subject from the standpoint of Lapicque's school, and makes an interesting attempt to give a quasi-physical basis to the tenets of that school upon the phenomena of curarization, summation, central inhibition and facilitation, and irreciprocal conduction. Particularly satisfying is it to find that the author is careful to collect the observations which oppose his theory as well as those which illustrate it. These objections are briefly discussed and plausible explanations are offered.

Thus Monnier's theory on the one hand brings under a single concept a very great mass of facts, and on the other it points to the exceptions as phenomena which demand further investigation both theoretically and by experiment.

W. A. H. R.

### *Free Radicals.* A General Discussion held by the Faraday Society, September, 1933. Pp. 248. 12*s.* 6*d.*

The present discussion covers a large amount of ground and various aspects of the subject receive treatment. Thirty separate communications, on many of which there were discussions, appear, so that it is impossible even to state the titles of all the papers contributed. The discussion was divided into two main sections: (i) free radicals of relatively long life, which included papers on the electronic structure and interaction of some simple radicals, the electronic levels of polyatomic molecules, and three dimensional models of the potential energy of triatomic systems; and (ii) free radicals of short life, subdivided into the chemical aspects, hydrocarbons, and the physical aspects, this part being subdivided into three sections. These sections included several subjects of interest to physicists such as free radicals produced in electric discharges and free radicals in spectroscopy. A paper on the application of molecular rays, however, is very brief and is mainly suggestive of future work. The use of positive rays in detecting free radicals, one of the oldest physical methods, is represented by a paper on the products of decomposition of some hydrocarbons, the chemical aspects of the thermal decomposition and combustion of hydrocarbons being also represented. Many of the papers deal with interesting theoretical aspects of the subject. This discussion, which was held at Cambridge, proved very interesting and fruitful and the resulting volume is one which contains many accounts of new work as well as critical discussions of results already known. The great advances which have been made in recent years in the study of the subject from many different points of origin are clearly apparent, and the volume is one of unusual interest and importance. Physicists who are interested in molecular structure and spectroscopy, in particular, will find in it much of interest. J. R. P.

*Scientific Inference*, by HAROLD JEFFREYS, F.R.S. (London: The Cambridge University Press.) 10*s*. 6*d*.

At the present time, when theories of measurement and the possibility (or impossibility) of the measurement of certain magnitudes is a matter of controversy, it is desirable to draw attention to Dr Jeffreys's valuable volume on scientific inference. The topics which he discusses are of vital interest to all physicists who occupy themselves with the philosophy of their subject, and his treatment of probability, sampling, quantitative laws, errors, physical magnitudes and mensuration is equally provocative and lucid.

His work is of wider range than this restricted catalogue would indicate; he is, in fact, primarily concerned with the principle that it is "possible to learn from experience and to make inferences from it beyond the data directly known by sensation," and, adopting this principle as a primitive postulate, has developed its consequences in various directions.

The argument is, as always with Dr Jeffreys, brilliantly clear, and is enlivened by the aptest and wittiest of illustrations. Who but the author would have seized on the immortal quatrain of Messrs Layton and Johnstone

"Oh, it ain't gonna rain no mo', no mo',
It ain't gonna rain no mo'!
How in the hell can the old folks tell?
'Tain't gonna rain no mo'!"

and employed it to head, most appropriately, a chapter on probability?

A. F.